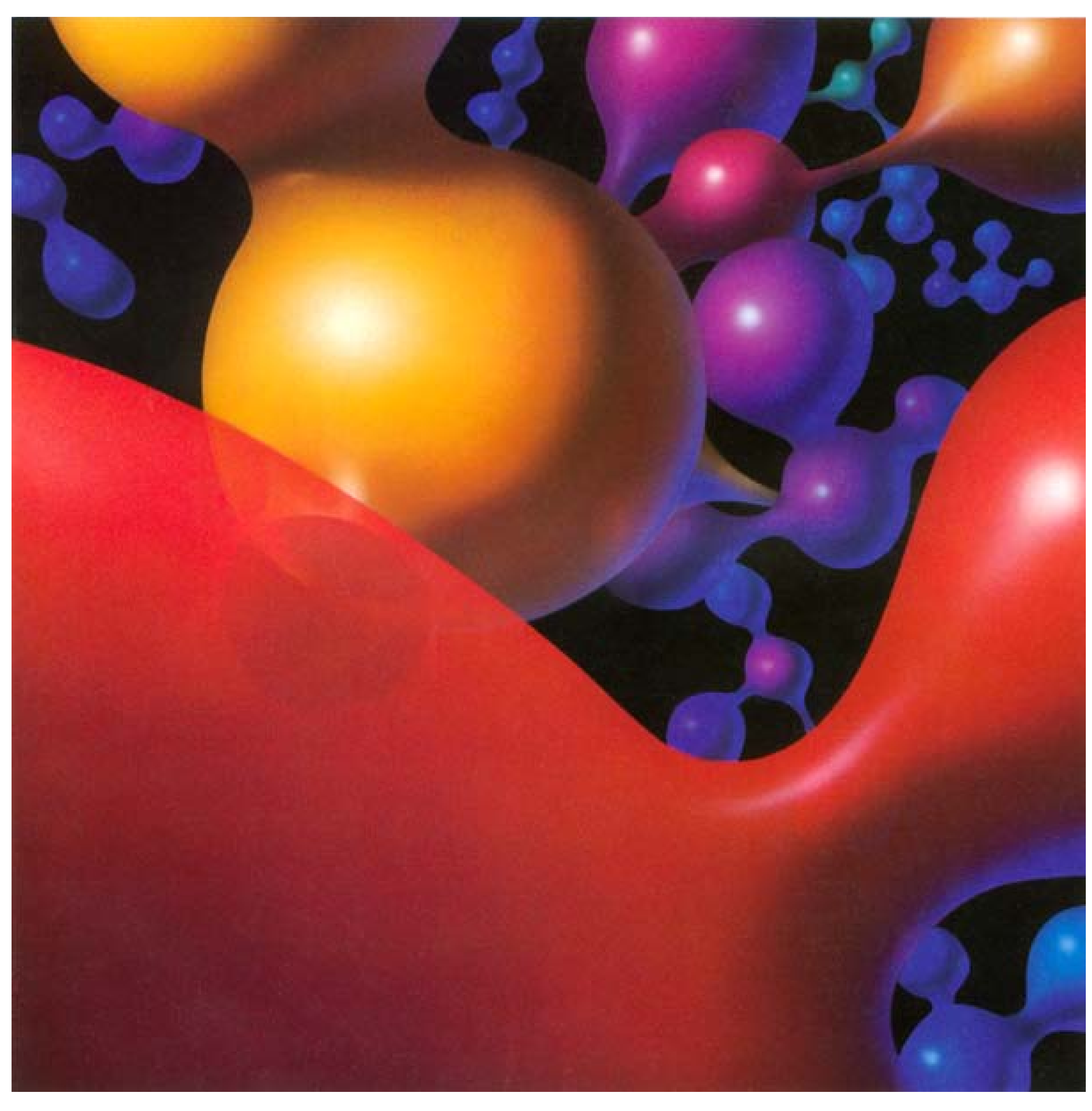
INVESTIGACION y CIENCIA

Edición española de SCIENTIFIC AMERICAN

PROTECCION DE REDES INFORMATICAS

M. C. ESCHER

LA BILIRRUBINA



NUEVOS MUNDOS

ENERO 1995
700 PTAS.



Enero de 1995 Número 220

Edición española de SCIENTIFIC AMERICAN 150 YEARS OF INNOVATION

SECCIONES

COLABORADORES DE ESTE NUMERO

Asesoramiento y traducción:

Esteban Santiago: *Epidemias de meningitis cerebroespinal*; J. García-Bellido: *El universo inflacionario autorregenerante*; Alfonso Susanna: *Claves genéticas del desarrollo floral*; Luis Bou: *Las metáforas de Escher*, *Redes informáticas seguras* y *Juegos matemáticos*; Ana M.ª Rubio: *¿Por qué hablan solos los niños?*; Josep Pla: *Una resolución de las paradojas de Zenón*; J. Vilardell: *Hace...*; Shigeko Suzuki: *De cerca*; José M. García de la Mora: *Cartas* y *Ensayo*

Ciencia y sociedad:

Juan P. Campos

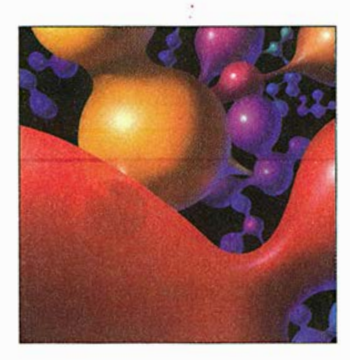
Ciencia y empresa:

M. Puigcerver, J.-D. Ros y Juan P. Campos

PROCEDENCIA DE LAS ILUSTRACIONES

Portada: Alfred T. Kamajian

Página	Fuente
8-9	Alan Reininger/ Contact Press Images
10	Dimitry Schidlovsky
11	Dana Burns-Pizer
12	Dimitry Schidlovsky
13	Johnny Johnson
14	Dimitry Schidlovsky
15	Patrick S. Moore
17-21	Servicio de Artes Visuales, Universidad de Stanford
22	Ian Worpole
23	Jared Schneidman/JSD
24-25	Elliot M. Meyerowitz
26	Patricia J. Wynne
27	*Fotografías, de arriba abajo:* G. N. Drews, cortesía de *Cell*; J. L. Bowman; George Retseck (*diagramas*)
28	J. L. Bowman (*arriba*), G. N. Drews, cortesía de *Cell* (*abajo*)
29	J. L. Bowman, cortesía de *Development* (*arriba*), T. Jack (*abajo*)
30	J. L. Bowman (*arriba*), D. Weigel (*abajo*)
44-49	M. C. Escher/ Cordon Art-Baarn, Holanda
50-51	Mark C. Flannery (*fotografía*)
52-53	Jana Brenning
54	Sam Ogden
57	R. Jonathan Rehg
59	Cortesía de Laura E. Berk
60	R. Jonathan Rehg
61	Lillian Holan, reproducida con permiso de HarperCollins Publishers, Inc.
63	Harold E. Edgerton 1992 Trust/ Palm Press, Inc.
64	Patricia j. Wynne
65-66	Johnny Johnson
67	Universidad de Oxford Museo de la Historia de la Ciencia
71-72	Javier González-Gallego, Claudio Tiribelli
73-75	Silvia Nuere
76-77	Javier González-Gallego, Claudio Tiribelli
85-87	Johnny Johnson
95	Bruce Hands/ Tony Stone Images



La figura de la portada describe una imagen insólita del cosmos: un mar burbujeante y ramificado de universos. Cada burbuja representa un universo, que crea otros universos, y así *ad infinitum*. Las leyes de la física de un universo particular, representadas por un color, tampoco están fijadas: un nacimiento podría producir una "mutación". Cada universo finalmente morirá, pero considerado en su globalidad el cosmos es eterno (véase "El universo inflacionario autoregenerante", por Andrei Linde, en este mismo número).

INVESTIGACION Y CIENCIA

DIRECTOR GENERAL Francisco Gracia Guillén
EDICIONES José María Valderas, *director*
ADMINISTRACIÓN Pilar Bronchal, *directora*
PRODUCCIÓN M.ª Cruz Iglesias Capón
Bernat Peso Infante
SECRETARÍA Purificación Mayoral Martínez
EDITA Prensa Científica, S. A. Muntaner, 339 pral. 1.ª – 08021 Barcelona (ESPAÑA)
Teléfono (93) 414 33 44 Telefax (93) 414 54 13

SCIENTIFIC AMERICAN

ACTING EDITOR Michelle Press
BOARD OF EDITORS John Rennie, *Associate Editor*; Marguerite Holloway, *News Editor*; Timothy M. Beardsley; W. Wayt Gibbs; John Horgan, *Senior Writer*; Kristin Leutwyler; Philip Morrison, *Book Editor*; Madhusree Mukerjee; Sasha Nemecek; Corey S. Powell; Ricki L. Rusting; Gary Stix; Paul Wallich; Philip M. Yam.
CHAIRMAN AND CHIEF EXECUTIVE OFFICER John J. Hanley
CO-CHAIRMAN Dr. Pierre Gerckens
DIRECTOR, ELECTRONIC PUBLISHING Martin Paul

SUSCRIPCIONES

Prensa Científica S. A.
Muntaner, 339 pral. 1.ª
08021 Barcelona (España)
Teléfono (93) 414 33 44
Fax (93) 414 54 13

Precios de suscripción, en pesetas:

	Un año	Dos años
España	7.700	14.000
Extranjero	8.600	15.800

Ejemplares sueltos:

Ordinario: 700 pesetas
Extraordinario: 900 pesetas

—Todos los precios indicados incluyen el IVA, cuando es aplicable.

— En Canarias, Ceuta y Melilla los precios incluyen el transporte aéreo.

— El precio de los ejemplares atrasados es el mismo que el de los actuales.

DISTRIBUCION

para España:

MIDESA

Carretera de Irún, km. 13,350
(Variante de Fuencarral)
28049 Madrid Tel. (91) 662 10 00

para los restantes países:

Prensa Científica, S. A.
Muntaner, 339 pral. 1.ª – 08021 Barcelona
Teléfono (93) 414 33 44

PUBLICIDAD

GM Publicidad
Francisca Martínez Soriano
Menorca, 8, bajo, centro, izquierda.
28009 Madrid
Tel. (91) 409 70 45 – Fax (91) 409 70 46

Cataluña y Baleares:
Miguel Munill
Muntaner, 339 pral. 1.ª
08021 Barcelona
Tel. (93) 321 21 14
Fax (93) 414 54 13

Difusión controlada

ISSN 0210136X Dep. legal: B. 38.999 - 76

Fotocomposición: Tecfa. Línea Fotocomposición, S.A. Almogàvers, 189 - 08018 Barcelona
Fotocromos reproducidos por Scan V2, S.A., Avda. Carrilet, 237 - 08907 l'Hospitalet (Barcelona)
Imprime Rotographik, S.A. Ctra. de Caldes, km 3,7 - Santa Perpètua de Mogoda (Barcelona)

Printed in Spain - Impreso en España

Los espacios en gris
corresponden a publicidad
en la edición impresa

Hace...

...cincuenta años

SCIENTIFIC AMERICAN: «El nuevo giróscopo Sperry para el control de la orientación proporciona a los pilotos de avión una indicación de la posición de su aparato, aun cuando no se vea la superficie terrestre. Este instrumento posibilita la ejecución de cualquier maniobra aerobática o acrobática sin la referencia visual de la superficie terrestre.»

«'En tiempos normales, de paz, la investigación es una actividad privada, no pública', afirma el Honorable Robert P. Patterson, subsecretario de Guerra. 'Lo que podría hacer el gobierno, si así se le pide, es facilitar información y apoyo económico. Podría brindar consejo, e incluso directrices. Pero en las investigaciones pacíficas normales, no debe asumir el control.'»

«El progreso científico, al igual que otros aspectos del esfuerzo humano, depende de la capacidad para comunicar ideas a los demás mediante el lenguaje. No se requiere del científico que sea un orador ni un hábil comunicador; basta con que se exprese claramente en lenguaje ordinario.»

«Tras casi dos años de desarrollo y ensayos de laboratorio, la Reynolds Metals Company comunicó recientemente disponer ya de 'la aleación de aluminio más resistente que hasta ahora se haya ofrecido para uso comercial'. La aleación, conocida como R303, contiene magnesio, zinc y cobre, y ofrece una resistencia a la compresión que triplica la del acero estructural.»

«Quizá la más interesante y prometedora de las aplicaciones propuestas para los plásticos reforzados con vidrio haya que buscarla en los modelos para elementos prefabricados de cocina y cuartos de baño de tamaño reducido y capacidad portante. Los dos conjuntos de dos lados, completados con medios de almacenamiento, se pretende que ocupen únicamente 0,65 metros cuadrados y sean capaces, no obstante, de soportar toda la estructura de una casa.»

«Son demasiadas las personas de edad que llevan el timón de la ciencia, y ésta necesita de la originalidad de la juventud para captar las posibilidades de servir a la humanidad.»

...cien años

SCIENTIFIC AMERICAN: «A pesar de la apatía casi general, un pequeño grupo de personas previsoras lleva años apremiando ante la necesidad de un sistema de gestión racional de los bosques de nuestro territorio nacional. Carecemos todavía de una política forestal sistematizada, pero al menos ya hay personas en puestos de responsabilidad que consideran que es éste un tema del que merece la pena hablar.»

«La rapidez en la transmisión de las noticias se ha convertido en una de las necesidades más imperiosas de nuestra época. Un nuevo telégrafo impresor permite reproducir a distancia el texto impreso por una máquina de escribir. El manuscrito a transmitir se reproduce simultáneamente en las estaciones receptoras instaladas en los domicilios de los distintos abonados.»

«La 'voz americana' disfruta de una reputación poco envidiable. Es proclive a ser chillona, estridente, sin modular. Estas cualidades agravan innecesariamente la vida social, ya que alteran los nervios y aumentan la propensión al agotamiento.»

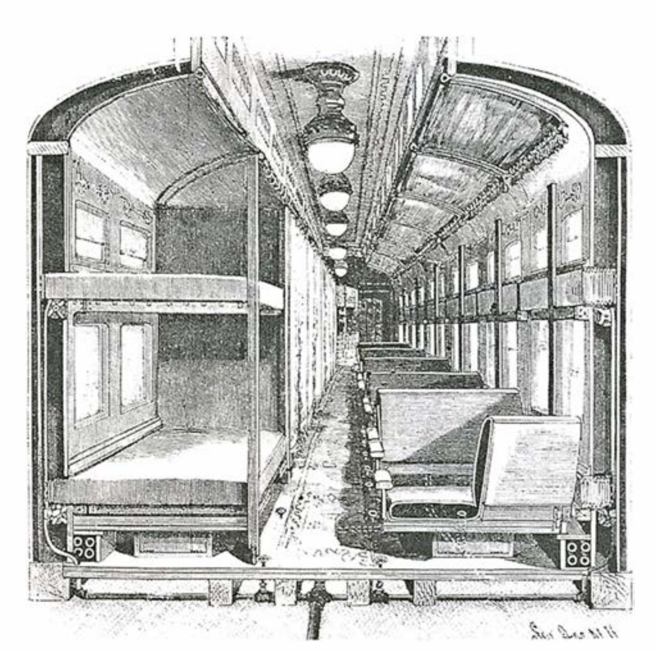
«En verdad, cada día se estrecha más el campo de utilización que antaño poseía el caballo. Al vapor, a la electricidad y a la ubicua bicicleta se les ha sumado el aliado del gas explosivo, tan ingeniosamente aplicado a la propulsión de vehículos que amenaza con desbancar al caballo por completo.»

«Actualmente existen seis millones de instalaciones agropecuarias en Estados Unidos. En ellas vive aproximadamente la mitad de la población de la república, unos treinta millones de personas, quienes generan más del 74 por ciento del total de las exportaciones del país.»

«La última de las modas higiénicas parisienses es usar vidrio poroso para las ventanas. Con éste, la luz penetra sin obstáculos y los poros dejan pasar el aire. Los minúsculos orificios son demasiado estrechos para facilitar corrientes de aire, mientras que brindan una saludable y continua ventilación por todo el piso.»

«Para ciertas cosas, el tamaño es una característica deseable; pero, en otras, la pequeñez es lo más deseable. En el caso de las baterías, la menor, más liviana y compacta de las que es posible fabricar produce una corriente intensa (2 ampères) con una tensión razonablemente alta (1,1 volts). Consiste en una célula de zinc, cerrada con un tapón de caucho duro, que contiene un electrodo formado de cloruro de plata fundido.»

«En la ilustración se muestra el coche combinado de salón y dormitorio, un notable invento registrado por Mr. Lindford F. Ruth. Los cojines y colchones de asientos y camas están conectados a las tuberías de aire comprimido del tren. Aquéllos no son sino cámaras estancas de caucho blando u otro material adecuado que pueden inflarse abriendo las válvulas de los tubos de conexión, o bien desinflarse y guardarse ocupando poco espacio.»



Coche salón-cama de Ruth

Cartas

Cálculo de π

Señor director:

El número de septiembre de *Investigación y Ciencia* incluye en su sección de Juegos de Ordenador un artículo de J. P. Delahaye en el que presenta un programa para el cálculo de 2400 cifras de π. Tan notable como el programa me ha parecido el comentario del autor: "No he conseguido averiguar en qué principios se basa la construcción de este extraordinario programa".

Incluso para un lego en estas materias como yo, no es difícil observar que las variables *d*, *b* y *g* del programa se usan para calcular expresiones de la forma $d = b/(2b+1)$.

Hay un conocido desarrollo:

$$\text{arctgz} = \frac{z}{1+z^2}\left[1 + \frac{2}{3}\frac{z^2}{1+z^2} + \frac{2\times 4}{3\times 5}\left(\frac{z^2}{1+z^2}\right) + \ldots\right]$$

que para z = 1 conduce a:

$$\frac{\pi}{2} = 1 + \frac{1}{3}\left(1 + \frac{2}{5}\left(1 + \frac{3}{7}(1 + \ldots)\right)\right)$$

y que debe ser la base de los cálculos.

Aparentemente, cada 14 términos de la serie contribuyen con cuatro dígitos al valor de π, y el programa utiliza esto de una manera extraordinariamente concisa y elegante para construir bloques de cuatro cifras usando aritmética modular exacta.

De lo que no estoy muy seguro es de la viabilidad del programa cuando se quiera calcular algunos millones de cifras, como se ha conseguido con otros procedimientos. El doble bucle que contiene, obliga a que el número de operaciones crezca como el cuadrado del número de cifras a calcular. Suponiendo que las 2400 cifras se calculen en diez segundos, el llegar hasta 2.400.000 supondría un millón de veces más, casi ciento dieciséis días.

Es probable que los programas basados en el uso de las propiedades de las funciones elípticas sean más rápidos y eficientes, —aunque dudo de que más elegantes—, pero esto es algo sobre lo que no tengo suficiente fundamento para opinar.

José T. Díez Roche
Catedrático de la Escuela Técnica Superior de Ingenieros Navales
Madrid

Mandíbula de Dmanisi

Señor director:

La publicación, en la sección de Ciencia y Sociedad del número de septiembre, de un suelto firmado por el profesor de Bonis me ha sumido en un mar de perplejidades. ¿Es tan poco lo que yo sabía de la paleontología humana de Europa? ¿He estado suministrando información incorrecta a mis alumnos durantes años? ¿Me habrán engañado mis ojos?

El origen de mis dudas resulta de una serie de curiosas aserciones contenidas en el escrito, algunas de las cuales son admisibles y otras no: por ejemplo, la presencia de Homínidos fósiles en Europa en series "mucho más antiguas" es admisible si utilizamos la descripción más amplia posible de la familia Hominidae, cosa en la que la mayor parte de los primatólogos no estaría de acuerdo, pero que no deja de ser un asunto de convenciones.

Igualmente convencional es el hecho de situar a Georgia en Europa. En realidad, sólo la relativamente reciente islamización de la península de Anatolia ha reforzado el sentimiento de que el Cáucaso separa dos continentes; basta recordar que, en la época grecorromana, era la orilla sur y no la norte del Mar Negro la que era culturalmente europea.

La afirmación más comprometida de entre las plausibles es la de que sería la primera vez que se habría asignado con certeza "una de las dos especies" (*Homo habilis* u *H. erectus*) al continente europeo. Cualquiera que tenga a mano los principales textos de paleoantropología de los años posteriores al de 1960 verá que, indefectiblemente, los restos de Mauer, Verteszöllös, Bilzingsleben y, con más reservas, los de Aragó eran sistemáticamente atribuidos a *Homo erectus*. Sólo recientemente, la obsesión de una importante corriente de paleontólogos por buscar un linaje plausible a los Neandertales ha dado en la solución de resucitar al *H. heidelbergensis* como antecesor de aquéllos.

Así, se ha desatado en algunos medios una campaña para convencernos de la existencia de un proceso de *neandertalización* a lo largo del Pleistoceno Medio europeo, campaña con algunos rasgos pintorescos, como la de cierto prestigioso paleontólogo que en 1984 no ve una fosa suprainíaca de tipo neandertal en el fósil de Swanscombe y sí la ve en 1988. En este tipo de campaña es donde hay que inscribir la afirmación de de Bonis de que Mauer presenta ciertas características de antepasado de neandertal: en realidad, la única característica que Mauer tiene en común con muchos neandertales (no todos) es el taurodontismo de los molares.

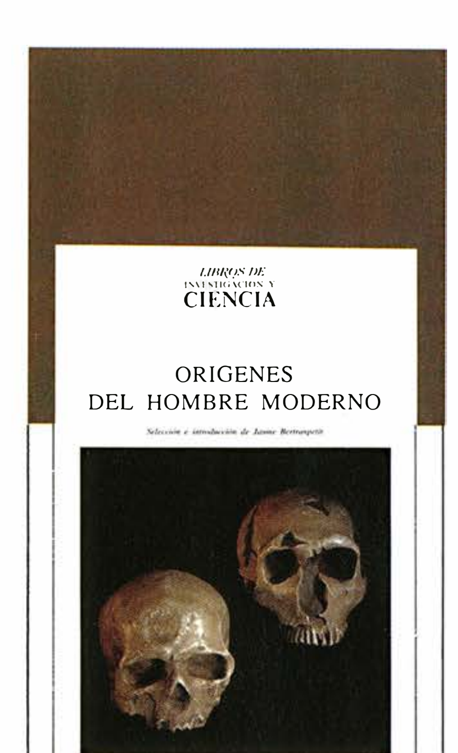
Pero lo importante es lo que sigue. Hasta ahora no hemos hecho más que discutir sobre apreciaciones y teorías, y aunque las del profesor de Bonis son extremadas, a mi modo de ver, no resultan radicalmente erróneas. Sin embargo, la afirmación de que "la mandíbula de Mauer se separa de la de Dmanisi por la presencia de un espacio retromolar notable", afirmación realzada por una capciosa elección de las ilustraciones, es, sencillamente, una barbaridad: basta orientar correctamente el resto de Mauer para observar que no tiene espacio retromolar.

D. I. Toja
Facultad de Biología
Universidad de Barcelona

Respuesta de Louis de Bonis

En respuesta a la carta de D. I. Toja sobre mi artículo "Paleoantropología europea: Raíz georgiana", me permito aportar algunos datos complementarios informativos. El primero, de poca monta, concierne a los límites de nuestro continente europeo.

Cierto que, en la medida en que ningún océano nos separa de Asia,



los límites entre Europa y el continente asiático pueden ser materia de discusión. Sin embargo, por regla general, las riberas septentrionales del Mar Negro son consideradas como europeas y, en una Europa que se extiende desde el Atlántico hasta los Montes Urales, puede ponerse a Georgia en los confines meridionales de nuestro continente.

Pero, entrando en un terreno más serio, por lo que atañe a la antigüedad en Europa de los homínidos, es evidente que el artículo toma este último término en un sentido restringido. Habiendo yo mismo descrito, en numerosas publicaciones, restos de *Hominidae* en el Mioceno superior de Grecia, sé muy bien que nuestra familia, en el sentido amplio que le dan la mayoría de los sistemáticos actuales, está presente en nuestras regiones desde hace muchísimo tiempo. Había, pues, que entender la palabra 'homínidos' como *hominianos*, grupo que abarca sólo a los australopitecos y al género *Homo*.

Ahora bien, la cuestión se circunscribía, inclusive, al mismo género *Homo* y el problema consistía en saber si la especie *Homo erectus* había sido registrada ya en Europa, según lo habría podido indicar la atribución de este nombre a ciertos restos como los de Vertesszöllös, Bilzingsleben o Tautavel, o si se trataba de un abuso de lenguaje para calificar unos fósiles ya orientados en una particular dirección evolutiva, concretamente en la que lleva hasta el hombre de Neandertal.

Si, como piensan muchos sistemáticos modernos y pienso yo con ellos, la sistemática debe conformarse tanto como sea posible a la filogenia, el nombre de *Homo erectus* se ha de reservar para los fósiles que precedan a la dicotomía entre los dos taxones, *Homo sapiens sapiens* por un lado y *Homo sapiens neanderthalensis* por el otro.

En Europa todos los fósiles humanos que preceden al hombre de Neandertal presentan, cuando son lo bastante completos, unos rasgos derivados característicos de los neandertalianos. El mejor ejemplo de ello es el cráneo de Petralona, descubierto en el norte de Grecia. La faz de este soberbio espécimen muestra unas estructuras anatómicas típicas de un hombre de Neandertal (ausencia de fosa canina, maxilar más desarrollado, posición muy baja del foramen infraorbitario...), mientras que la región occipital conserva un matiz primitivo próximo al de la que se conoce como propia del *Homo erectus*. Otros especímenes, aunque menos completos, presentan también rasgos neandertalianos ya sea en la faz (Tautavel) ya en la parte posterior del cráneo (Swanscombe o Biache-Saint Vaast). En lo concerniente a restos más fragmentarios, como los de Vertesszöllös y Bilzingleben, no pueden verse en ellos más que algunas características primitivas (reborde suborbitario en uno o toro occipital en los dos), pero muy probablemente unas piezas más completas nos mostrarían algunos rasgos neandertalianos. Por lo demás, cada uno de los fragmentos de estos dos yacimientos es afín a su homólogo del cráneo de Petralona.

Todos estos fósiles son, pues, posteriores a la separación de los dos linajes, del hombre moderno y del hombre de Neandertal, y también es este el caso de la mandíbula de Mauer, considerada hasta hoy como el más antiguo documento de la presencia humana en Europa. Las características neandertalianas no son muy marcadas, pero en la mandíbula se ve por lo general el taurodontismo, esto es ya indiscutible, y el espacio retromolar. Este último aparece no sólo en una vista de tres cuartos por delante como la del artículo incriminado, sino también —aunque es menos importante que el de ciertos neandertalenses— en una vista de la parte interior de la mandíbula de Mauer a propósito de la cual E. Genet-Varcin escribía: "...su aspecto general la aproxima a la mandíbula de los neandertalenses" y se distingue de las mandíbulas de *Homo erectus* porque "los dos bordes de su cuerpo, el alveolar y el inferior, se acercan más el uno al otro yendo hacia M/3... la rama montante es netamente menos alta aun siendo también larga. Esta mandíbula prueba que las formas neandertaloides existieron en Europa desde tiempos muy remotos."

Ahora bien, parece ser que la mandíbula de Dmanisi no presenta —hasta más amplio informe— ninguno de los rasgos neandertalianos característicos. Desde el punto de vista anatómico podría, pues, representar un estadio anterior a la dicotomía entre las dos subespecies de *Homo sapiens* y merecer el nombre de *Homo erectus*.

Cuando se está tratando de ampliar la Unión Europea, no hay que restringir demasiado los límites de nuestro continente y conviene admitir que el europeo más antiguo pudo muy bien habitar en Georgia.

LOUIS DE BONIS
Universidad de Poitiers

Epidemias de meningitis cerebroespinal

Enfermedad debilitante y a menudo mortal, la meningitis sigue siendo común en muchos países en vías de desarrollo. Los nuevos descubrimientos podrían permitirnos, muy pronto, prever y controlar sus brotes epidémicos

Patrick S. Moore y Claire V. Broome

A mediados de abril de 1988 la epidemia de meningitis declarada en Ndjamena, capital del Chad, estaba en pleno apogeo. Había comenzado, mediado febrero, con unos pocos pacientes; al cabo de cuatro semanas, ingresaban unos 150 diarios en el Hospital Central de la ciudad. Ocupadas todas las camas disponibles, hubo que atender a los enfermos en tiendas de campaña del ejército instaladas en el recinto asistencial. Pese a los esfuerzos del Ministerio de Sanidad y de organizaciones voluntarias extranjeras, la epidemia se propagó. La escasez de medicamentos hizo más duro el trabajo del personal sanitario que resistió días de trabajo agotador. Aunque se emprendió una campaña de vacunación en masa para atajar la epidemia, se temía por momentos la quiebra del endeble sistema sanitario del país.

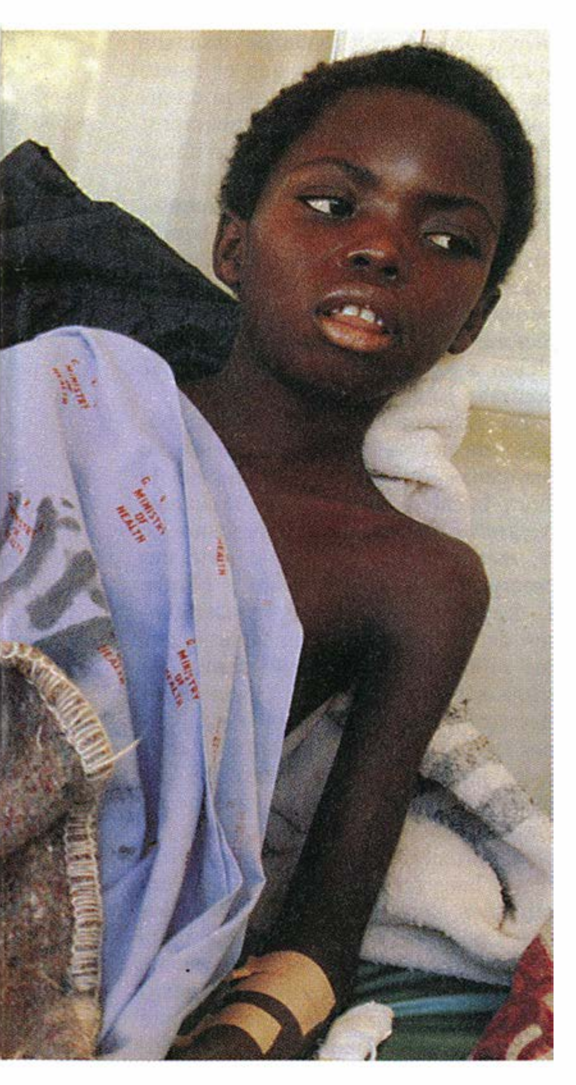
Para cuando terminó el azote, 4500 personas, según las estadísticas oficiales, habían contraído la meningitis. Pero hubo otros centenares, o



incluso millares, sin registrar. En el Chad, al igual que en muchos países africanos, la atención médica queda fuera del alcance de quienes residen lejos de las principales ciudades. A los enfermos que vivían a más de un día de camino de un puesto sanitario no les llegaba el tratamiento antibiótico. Muchos murieron o les quedaron lesiones cerebrales permanentes.

El caos y la miseria producidos por la epidemia en el Chad son característicos de la mayoría de los brotes de meningitis meningocócica. Esta enfermedad se instaura con vertiginosa rapidez. Una persona sana desarrolla primero fiebre y una sensación de debilidad similar al estado griposo. En cuestión de horas, tales síntomas evolucionan hacia una intensa cefalea, rigidez de nuca y fotofobia. Si no se trata, el paciente entra en coma y, por fin, en choque terminal. Aunque raras ahora en los países adelantados, las epidemias siguen afectando a las naciones en vías de desarrollo; en pocas semanas el flagelo puede aterrorizar a todo un país.

¿Por qué se producen esas epidemias? ¿Qué induce a una enfermedad como la meningitis a vivir silente incrustada en una población durante años y brotar de forma súbita? Aunque muchos misterios rodean todavía a esta enfermedad de capacidad letal, su peculiar epidemiología ofrece algunas pistas sobre las causas desencadenantes y su prevención. La enfermedad comprende ciclos de incidencia que pueden corresponder a cambios ambientales, a patrones poco habituales de inmunidad e incluso a una asociación con otras enfermedades infecciosas. La investigación clínica y la aplicación de nuevas técnicas biológicas han sacado a la luz algunos de sus ominosos secretos.



La bacteria que causa la meningitis meningocócica se llama *Neisseria meningitidis* o meningococo, un pariente cercano de la bacteria de la gonorrea. A diferencia de la responsable de esta última, los meningococos suelen colonizar el revestimiento de la faringe y desde allí se propagan a través de las secreciones respiratorias. El organismo es tan común, que podría considerarse parte integrante de la flora bucal del hombre: entre un 2 y un 10 por ciento de las personas sanas puede ser portador en un momento dado de meningococos.

El revestimiento epitelial de la garganta actúa de barrera natural al paso de bacterias, aunque en ocasiones se rompe el equilibrio entre colonización e invasión. Este desequilibrio desemboca en una enfermedad grave. La meningitis meningocócica comienza cuando el microorganismo invade el torrente sanguíneo y atraviesa las meninges —membranas que rodean el cerebro y la médula espinal— para acceder al líquido cefalorraquídeo que baña el sistema nervioso central y actúa a modo de medio de cultivo para el rápido crecimiento de las bacterias, que después inflaman el revestimiento meníngeo.

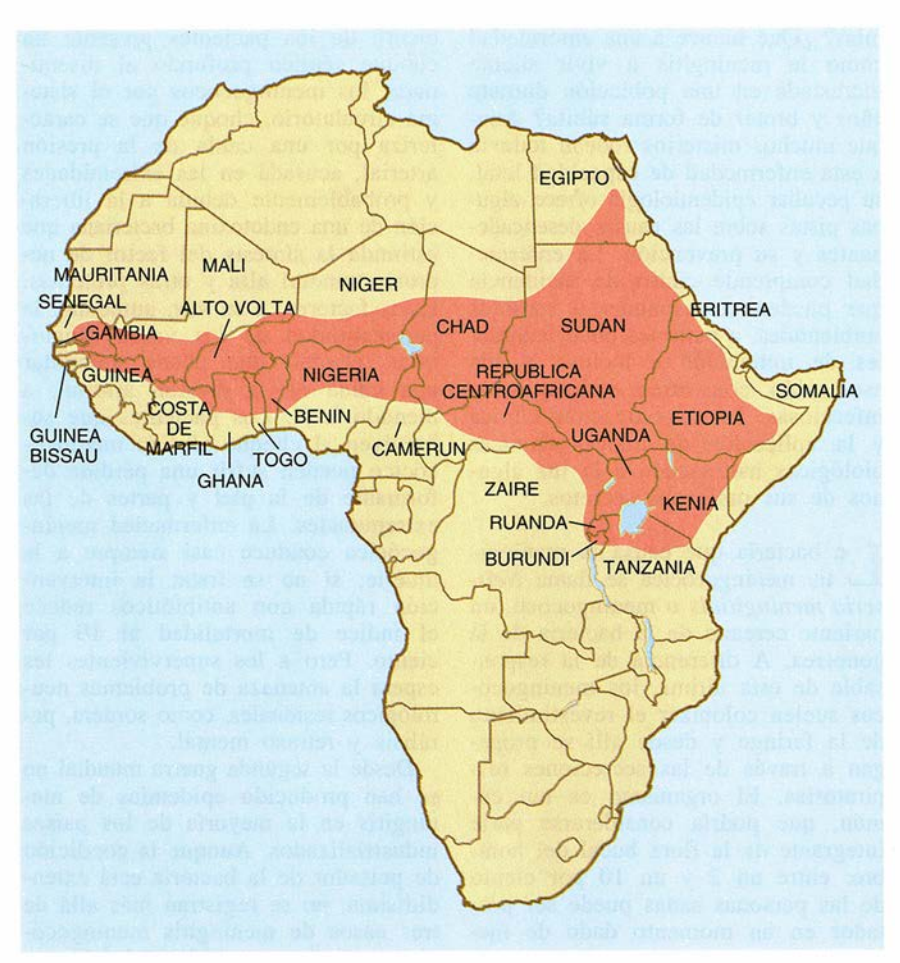
Los síntomas típicos, entre los que se cuentan fiebre, rigidez de nuca, cefalea y coma, son la consecuencia de esta inflamación. Hasta un 30 por ciento de los pacientes presenta un choque séptico profundo al diseminarse los meningococos por el sistema circulatorio, choque que se caracteriza por una caída de la presión arterial, acusada en las extremidades y probablemente debida a la liberación de una endotoxina bacteriana que estimula la síntesis del factor de necrosis tumoral alfa y otras proteínas. Estos factores, a su vez, aumentan la permeabilidad de los vasos sanguíneos, cambio que puede precipitar una caída de la presión arterial, a menudo letal. Los pacientes que sobreviven al choque séptico meningocócico pueden sufrir una pérdida deformante de la piel y partes de las extremidades. La enfermedad meningocócica conduce casi siempre a la muerte, si no se trata; la intervención rápida con antibióticos reduce el índice de mortalidad al 10 por ciento. Pero a los supervivientes les espera la amenaza de problemas neurológicos residuales, como sordera, parálisis y retraso mental.

Desde la segunda guerra mundial no se han producido epidemias de meningitis en la mayoría de los países industrializados. Aunque la condición de portador de la bacteria está extendidísima, no se registran más allá de tres casos de meningitis meningocócica endémica por 100.000 habitantes en Estados Unidos. Pero en el Tercer Mundo la situación es otra. Hubo más de 40.000 casos documentados durante la epidemia que asoló a Etiopía en 1989 y es posible que en el curso de los años sesenta se vieran afectadas hasta tres millones de personas. La súbita afluencia de centenares y millares de casos colapsa los sistemas sanitarios, a menudo rudimentarios, de las naciones pobres.

Por su carácter esporádico e impredecible, el proceso epidémico resulta difícil de abordar. Los epidemiólogos llaman "cinturón de la meningitis" a la ancha banda de países situados al sur del Sahara, desde Gambia, en el oeste, hasta Etiopía. En esa región de extensas sabanas que cruza el continente africano, las epidemias son especialmente frecuentes y tienden a

1. PACIENTE CON MENINGITIS que descansa junto a su padre en un hospital africano. El tratamiento con antibióticos reduce el índice de mortalidad de esta enfermedad hasta un 10 por ciento. Pero la mayoría de la gente que enferma vive en áreas rurales de países en vías de desarrollo, sin acceso al tratamiento médico ni a las instalaciones hospitalarias.

PATRICK S. MOORE y CLAIRE V. BROOME son especialistas en salud pública y en prevención de enfermedades infecciosas. Moore enseña en la escuela de salud pública de la Universidad de Columbia. Broome es subdirectora del Centro de Control y Prevención de Enfermedades.



2. CINTURON DE LA MENINGITIS. Atraviesa el Africa central, desde Eritrea en el este hasta Gambia en el oeste. La gente que vive en esta región parece ser extraordinariamente sensible a las repetidas epidemias de meningitis. Los brotes declarados en el cinturón son consecuencia, la mayoría de las veces, de la infección por cierta cepa de bacterias, la correspondiente a los meningococos del serogrupo A.

sobre el ganado. La reducción de esa inmunidad de la comunidad podría ser responsable en parte de los modelos cíclicos que sigue la meningitis en Africa.

recurrir cada cinco a doce años. Cada oleada dura varios años.

En el curso de un año cualquiera, la meningitis sigue un segundo ciclo, éste anual: los casos alcanzan su máximo durante la estación seca y desaparecen con la llegada de las lluvias. Incluso durante el apogeo de una epidemia, la incidencia disminuye hasta niveles basales en la estación lluviosa, para volver a repuntar con la siguiente estación seca. De lo que se infiere que, cuando la bacteria está implantada en una población sensible, hay algún fenómeno de estación seca que determina la eclosión o no de la epidemia.

Estos rasgos misteriosos de la plaga meningocócica han venido intrigando a los epidemiólogos desde hace años. A diferencia de la enfermedad endémica, los factores de riesgo afectan a toda la población, no a individuos dispersos; los factores que determinan su inicio deben ser variables, pues varía con el tiempo el riesgo epidémico. En Estados Unidos, las personas con alto riesgo de contraer la meningitis meningocócica son aquellas que nacen con cierta deficiencia genética en su sistema del complemento (serie de proteínas sanguíneas activadas por los anticuerpos para destruir bacterias). Siendo muy rara la enfermedad en Norteamérica, cabría aceptar que una proporción sustancial de todos los casos deberíase a dicha alteración genética. Ahora bien, el número de individuos con esa predisposición tiende a permanecer constante; por tanto, es improbable que se trate de una causa significativa de una epidemia. De hecho, los estudios realizados en Nigeria y en Gambia han confirmado que no era frecuente encontrar pacientes con deficiencias del complemento durante esos brotes.

Pero otros factores del huésped sí podrían cambiar con el tiempo en una población; por ejemplo, sus niveles de anticuerpos contra los meningococos. En una población, el nivel general de inmunidad contra un microorganismo se denomina inmunidad del rebaño, expresión que atestigua su origen en las investigaciones

A finales de los sesenta, el grupo de Irving Goldshneider y Emil C. Gotschlich puso de manifiesto la importancia de las defensas del huésped contra los meningococos. Antes de la introducción de las vacunas, en la década siguiente, los reclutas militares de Walter Reed eran especialmente sensibles a la meningitis. El grupo extrajo sangre de millares de soldados al comienzo del período de instrucción e hizo un seguimiento durante su estancia en el campamento. Conforme enfermaban, se analizaba el suero almacenado para determinar su capacidad de destrucción de meningococos y se comparaba con la actividad antimeningocócica de los sueros de sus compañeros sanos.

La enfermedad se presentaba fundamentalmente en los reclutas cuya actividad antimeningocócica sérica antes de caer enfermos era baja. Puesto que la mayoría de los adultos tiene anticuerpos protectores frente a los meningococos, este hallazgo resultaba sorprendente. Cierto es que el sujeto ha de hallarse expuesto al microorganismo para desarrollar anticuerpos. Pero la investigación con los reclutas manifestaba que los individuos sin anticuerpos y expuestos a los meningococos tenían mayores probabilidades de enfermar. No acababa de entenderse por qué la mayoría de las personas desarrollan anticuerpos protectores después de su primera exposición a los meningococos en vez de contraer la enfermedad.

La respuesta a esta paradoja podría encerrarse en las bacterias no patógenas que forman parte de la flora bucal normal de los humanos. *Neisseria lactamica*, un pariente de *N. meningitidis*, es uno de esos microorganismos. Ronald Gold y Martha L. Lepow, junto con los investigadores de Walter Reed, demostraron que los niños que sufren infecciones faríngeas producidas por *N. lactamica* desarrollan con frecuencia anticuerpos que son también protectores frente a los meningococos. La infección por un clase de bacterias no patógenas parece proteger contra la invasión por otra perteneciente a una cepa virulenta. Este fenómeno ofrece una explicación para el caso de los reclutas. Cualquier soldado que careciera de anticuerpos de protección cruzada desarrollada tras una infección infantil por *Neisseria* sería especialmente sensible porque las cepas meningocó-

cicas se propagan con rapidez entre los compañeros de barracón.

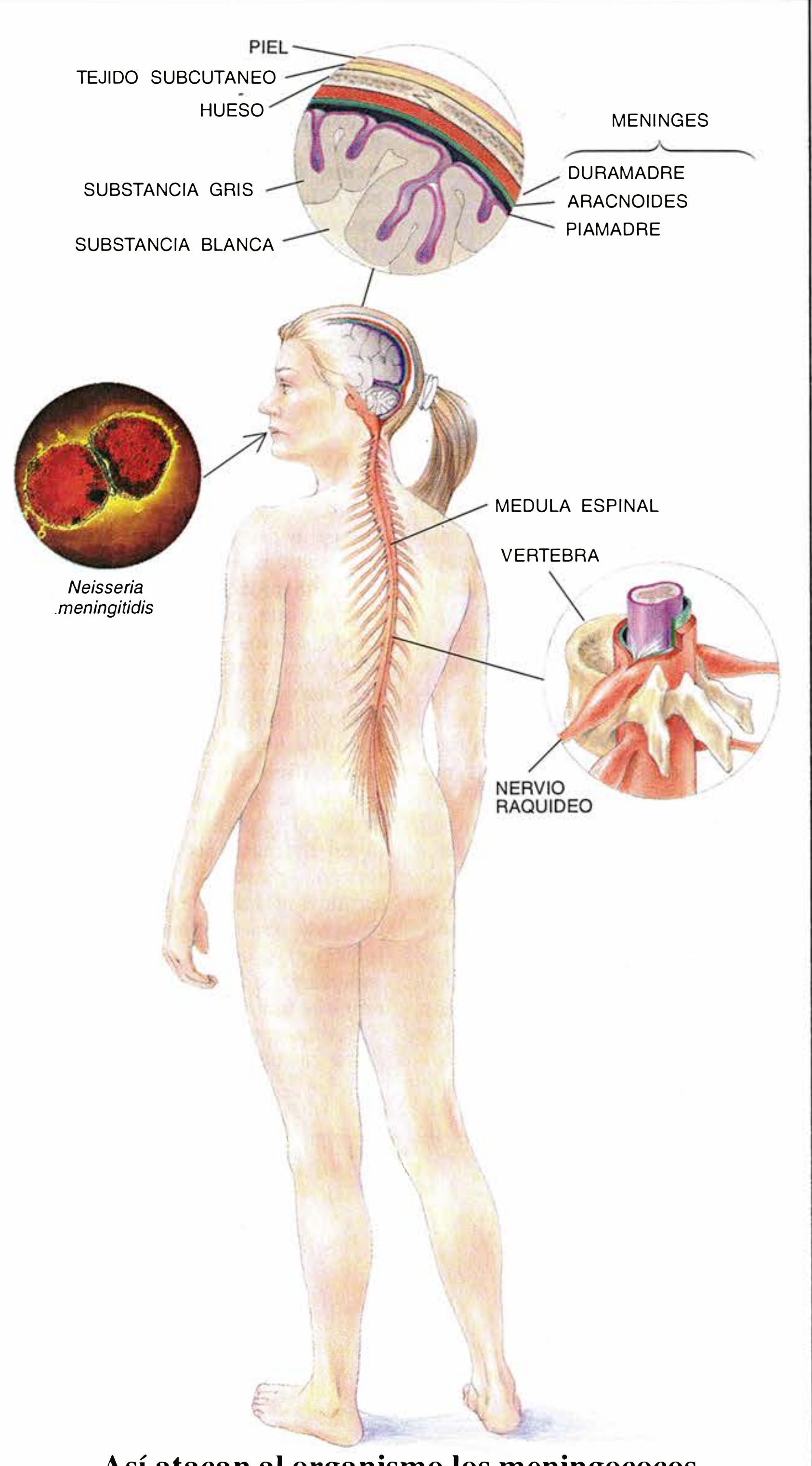
Se está investigando qué componentes de los meningococos producen la respuesta inmunitaria protectora. La cápsula de polisacáridos que rodea al microorganismo es un antígeno meningocócico —es decir, una molécula que estimula una respuesta inmunitaria. Las cepas meningocócicas poseen diferentes polisacáridos de naturaleza antigénica, de los que se han encontrado al menos 13 serogrupos, o tipos distintos. *N. meningitidis* del serogrupo A es la responsable de las epidemias masivas que periódicamente afectan a Africa, China e Hispanoamérica.

Es menos probable que otros serogrupos desencadenen epidemias, aunque se trate, en buena medida, de los responsables de la enfermedad endémica de Estados Unidos. Las vacunas producidas a partir de un polisacárido particular son muy eficaces contra el correspondiente serogrupo, pero no ofrecen protección cruzada. Por desgracia, el polisacárido del serogrupo B, que es el más común en Estados Unidos, no induce una respuesta inmunitaria prolongada.

De acuerdo con los trabajos del grupo de J. McLeod Griffiss y Robert E. Mandrell, así como los del equipo de Wendell D. Zollinger, habría, además de los polisacáridos, otros antígenos que desempeñen un papel importante en la inmunidad. Al no ser protección cruzada la desarrollada contra el polisacárido, se supone que la amplia inmunidad surgida de la infección por *N. lactamica* vendría de otros antígenos. En efecto, es muy posible que *N. lactamica* carezca de polisacárido capsular. Quizás otros componentes celulares —como las proteínas de la membrana externa y los lipooligosacáridos unidos a la membrana comunes a *N. lactamica y N. meningitidis*— confieran inmunidad.

En ese mecanismo inmunitario podría yacer la razón de los intervalos, de 5 a 12 años, que transcurren entre dos brotes epidémicos en Africa: tasas altas de infección meningocócica durante una epidemia provocarían una inmunidad natural generalizada, que después protegería a la población durante varios años; conforme decayera la inmunidad con el nacimiento de niños sensibles y con la pérdida natural de anticuerpos, la población se tornaría de nuevo vulnerable a la epidemia.

Con todo, no podemos atribuir exclusivamente a la pérdida de inmunidad el decurso sorprendente de la enfermedad en el "cinturón de la me-



Así atacan al organismo los meningococos

La bacteria causante de la meningitis cerebroespinal, *Neisseria meningitidis*, prolifera en la mucosa faríngea. Se produce la enfermedad cuando el organismo penetra en la sangre y llega a las meninges, membranas que recubren el cerebro y la médula espinal. El microorganismo crece con rapidez en ese entorno, inflamando las meninges y produciendo fiebre, rigidez de nuca, cefalea y, a menudo, coma. En un 30 por ciento de los pacientes, la bacteria libera una endotoxina que aumenta la permeabilidad de los vasos sanguíneos. En virtud de ello, se produce una caída de la presión arterial; el choque séptico consecutivo puede producir pérdida de piel y de zonas de las extremidades.

ningitis". Su carácter estacional denuncia la participación de factores ambientales. Buena parte de cuanto sabemos a propósito de éstos lo hemos aprendido en las investigaciones realizadas en Gambia y Nigeria por el equipo de Brian M. Greenwood. Este grupo observó que la transmisión de la enfermedad meningocócica procedía a lo largo de todo el año, aunque se manifieste sólo durante la estación seca.

Detectaron, además, niveles altos de anticuerpos en algunos poblados después de la estación de lluvias (sin que se registrara ningún paciente); ello sugería que los meningococos se habían propagado entre la población elevando la inmunidad sin causar enfermedad. De lo que se infiere que el carácter estacional de la meningitis en Africa no resulta de una mayor transmisión durante la estación seca. La desecación de la mucosa faríngea durante la estación seca podría aumentar la colonización meningocócica de los tejidos subyacentes.

Además del clima, las infecciones víricas de las vías respiratorias superiores podrían interesar las mucosas faríngeas, redoblando su vulnerabilidad a la invasión. La neumonía bacteriana, por ejemplo, puede asentarse tras una infección vírica. Nuestro grupo, perteneciente a la Sección de Meningitis y Agentes Patógenos Especiales de los Centros de Control y Prevención de Enfermedades (CDC), abordó esta posibilidad durante una serie de epidemias que comenzaron a mediados de los ochenta.

En agosto de 1987 las autoridades sanitarias de Nueva York nos informaron de dos viajeros que habían contraído meningitis en vuelos distintos, de regreso de Arabia Saudita. Durante la semana anterior, habían llegado noticias a la CDC de la aparición de meningitis meningocócica entre los participantes en la peregrinación musulmana anual a La Meca. Sin embargo, no se percibía con nitidez que se tratara del comienzo de una epidemia. (El brote de 1987 originado en La Meca produjo, por lo menos, 10.000 casos de meningitis.) Lee Haarrison, Gloria W. Ajello y uno de nosotros (Moore) se dirigieron al aeropuerto internacional John F. Kennedy para recibir a los aviones llenos de peregrinos.

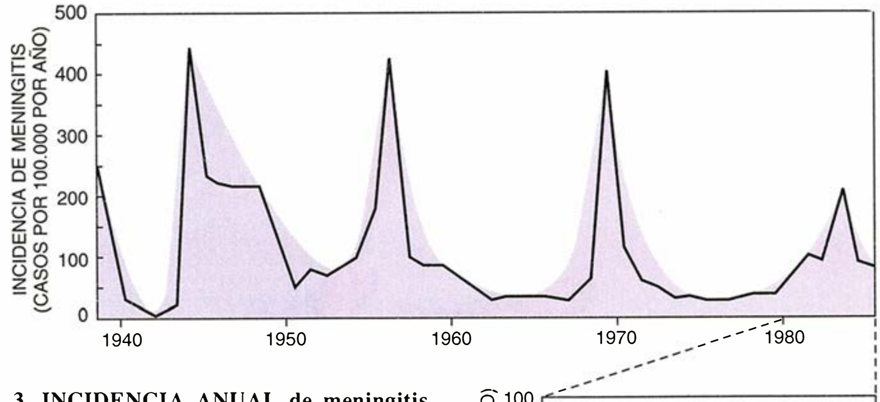
Instalamos un dispensario provisional en el vestíbulo del aeropuerto y administramos antibióticos como tratamiento profiláctico. De los 550 pasajeros que examinamos, el 11 por ciento de los que regresaban de La Meca era portador de meningococos del serogrupo A, cepa rarísima en Estados Unidos. Entre los portadores abundaban más los síntomas de resfriado (fiebre y dolores de garganta) que entre los no portadores. Pero, como sólo pudimos examinar portadores de meningococos y no pacientes con meningitis, estos datos son circunstanciales. El siguiente paso obligado era la búsqueda directa de infecciones en las vías respiratorias superiores de los pacientes con meningitis meningocócica.

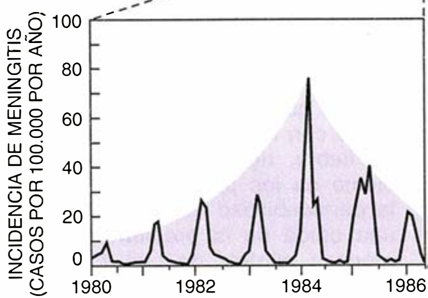
Pudimos entrar en esa cuestión durante la epidemia del Chad. Cierta mañana del mes de abril de 1988 recibimos una llamada de Theo Lippeveld. Nos avisó de que en Ndjamena, ciudad de 500.000 habitantes situada en el centro del "cinturón de la meningitis", se había desencadenado una grave epidemia meningocócica del serogrupo A. El Ministerio de Sanidad, bajo la dirección de P. Matchock Yankalbé, estaba organizando las medidas de control conjuntamente con la Bioforce francesa (un grupo de médicos dependiente del gobierno galo) y el Instituto Merieux de Marsella. Las autoridades del Chad nos concedieron permiso para abrir una investigación, y así partió para la capital un equipo del CDC que habría de trabajar con los médicos de su Hospital Central.

Al objeto de detectar infecciones respiratorias, formamos dos grupos equiparables en edad, sexo y barrio de residencia; uno, compuesto por personas con meningitis, y el otro, el grupo control, por personas sanas. Se recogieron los lavados nasales y se enviaron a John Hierholzer del CDC, en quien recayó la pesada tarea de procesar y cultivar centenares de lavados nasales. Los resultados fueron sorprendentes. En conjunto, los pacientes de meningitis tenían una probabilidad 23 veces mayor de portar un agente patógeno en las vías respiratorias superiores que sus controles correspondientes. No sólo era más probable en ellos la invasión de esos virus, sino que una gran proporción estaba infectada también por *Mycoplasma hominis*, una bacteria intracelular.

Estos hallazgos aducen una razón más del carácter estacional de las epidemias de meningitis: presumiblemente, la combinación de baja humedad e infecciones respiratorias coloca a la población en situación de riesgo. Hay que ahondar en la investigación y desentrañar el mecanismo en virtud del cual las infecciones respiratorias interaccionan con los meningococos. De todas formas, David S. Stephens y Zell A. McGee han demostrado ya en el laboratorio que, en circunstancias semejantes, las células del revestimiento faríngeo captan meningococos. Las infecciones respiratorias podrían estimular el proceso de captación; y cabe también que las infecciones respiratorias dañen directamente las mucosas o bloqueen allí a las células inmunitarias.

A mayor abundamiento, se han asociado infecciones respiratorias con la enfermedad meningocócica en los países industrializados, lo que es compatible con el hecho de que la enfermedad sea más frecuente en los meses de invierno, cuando proliferan





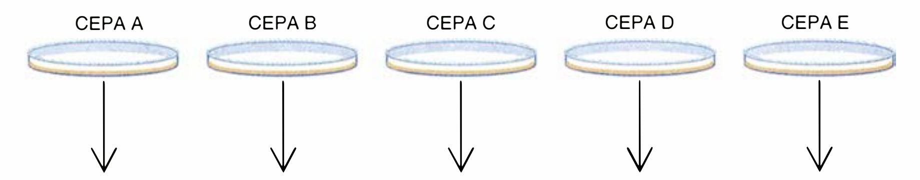
3. INCIDENCIA ANUAL de meningitis en Alto Volta (Burkina Faso). Sigue el patrón típico de los países del cinturón de la meningitis. Las epidemias, de varios años de duración, siguen un modelo crescendo-decrescendo (*arriba*). Sin embargo, incluso durante una epidemia la frecuencia de casos tiene un carácter estacional. Como se ve en el otro gráfico más detallado (*derecha*), los brotes se producen sólo durante la estación seca, desde enero hasta junio.

los virus del resfriado. Keith A. V. Cartwright, Dennis M. Jones y James M. Stuart señalaron no hace mucho una asociación semejante entre enfermedad meningocócica y gripe. La misma relación la encontró el equipo de Bruno Hubert. Esta línea de investigación puede llevar al descubrimiento de nuevas vías para predecir la aparición de un brote epidémico en esos países. Mas, por paradójico que parezca, seguimos sin comprender las razones del carácter tan estacional de la infecciones de las vías respiratorias superiores.

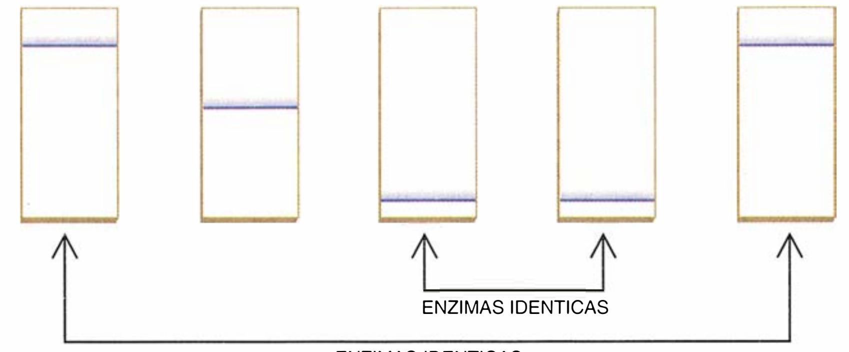
Mientras llevábamos a cabo nuestro estudio en el Chad, el Ministerio de Sanidad controló la epidemia; se detuvo su curso con la intervención del grupo de Harvard, médicos del Hospital Central de la capital y voluntarios extranjeros; numerosos gobiernos occidentales se volcaron en su ayuda. Casi el uno por ciento de la población de Ndjamena contrajo la meningitis (en el grupo de escolares y el de soldados, la cifra alcanzó quizás el 10 por ciento). La incidencia hubiera sido aún mayor si no se hubieran tomado las medidas adecuadas de control. Fue un ejemplo claro de la movilización internacional necesaria cuando brota una epidemia.

Si los anticuerpos, el clima y las infecciones respiratorias son factores importantes para que se perfile una epidemia, ¿qué papel desempeña el propio organismo? El nuevo campo de la epidemiología molecular puede ayudarnos a encontrar la respuesta. Incorporando técnicas de biología molecular —secuenciación del ADN y electroforesis de enzimas— los epidemiólogos pueden desentrañar ahora los misterios que rodean a los brotes producidos por los descendientes de una cepa única, o clon. Estas técnicas han permitido ya seguir la pista de diversos virus y bacterias patógenos y han demostrado su eficacia a la hora de descubrir un caso de transmisión del virus de la inmunodeficiencia humana de un dentista de Florida a varios de sus pacientes.

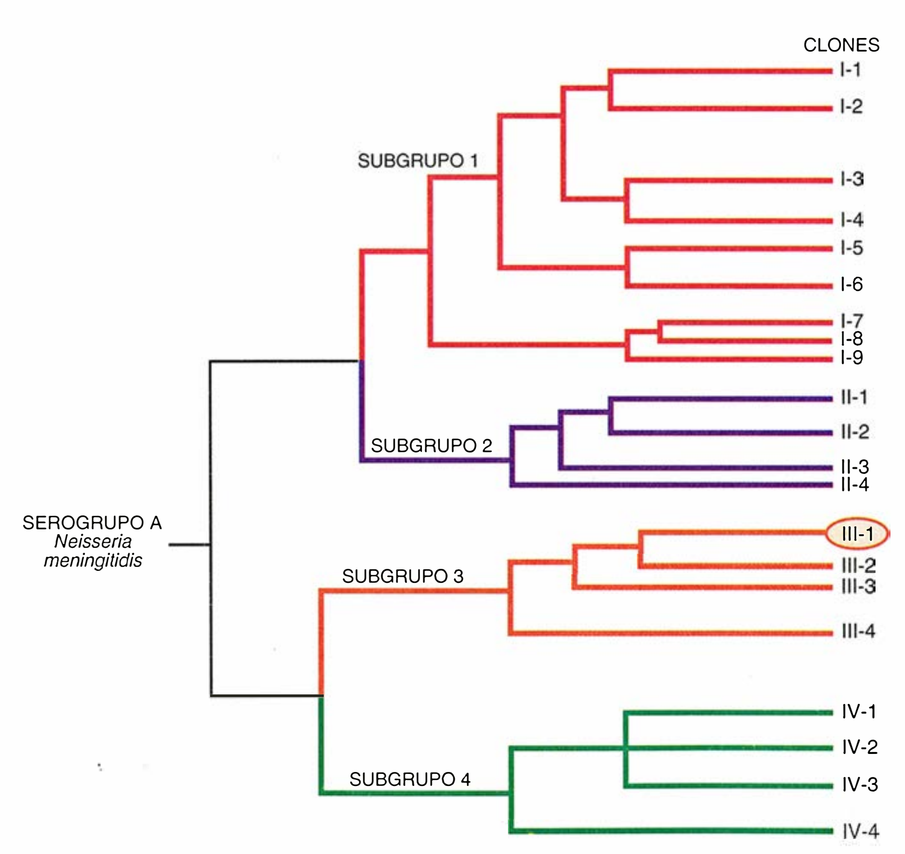
Exactamente igual que los biólogos pueden hacer uso de la acumulación de mutaciones a lo largo del tiempo para rastrear la divergencia evolutiva entre dos especies, los epidemiólogos pueden analizar el ADN, buscando mutaciones, para distinguir dos cepas del mismo microorganismo en su tránsito de unas poblaciones humanas a otras. Para localizar al meningococo del serogrupo A se ha utilizado muy a menudo la electroforesis de enzimas de múltiples loci, un método indirecto para examinar



SE APLICA UNA CORRIENTE ELECTRICA AL GEL QUE CONTIENE LAS ENZIMAS; LAS QUE SON IDENTICAS SE DESPLAZAN HASTA EL MISMO LUGAR EN EL GEL



4. ELECTROFORESIS ENZIMATICA. A ella se acude para detectar divergencias entre cepas bacterianas. Las enzimas intracelulares de varias cepas se colocan en un gel y se separan cuando se aplica a la mezcla una corriente eléctrica. Cada enzima se identifica por mutaciones individuales, que marcan su naturaleza. Las enzimas de cepas diferentes migrarán a lugares distintos del gel. Podemos analizar así múltiples enzimas para determinar si son o no idénticas.



FUENTE: Mark Achtman, Instituto Max Planck de Genética Molecular, Berlín

5. MENINGOCOCOS DEL GRUPO A, antaño considerados homogéneos, pertenecen al menos a 21 líneas o clones, según revela la electroforesis enzimática. Uno de esos clones, el III-1, ha causado recientemente epidemias en Asia, Oriente Medio y Africa.

el parentesco genético entre cepas diferentes.

Esta técnica poderosa, de apariencia sencilla, se basa en la detección de mutaciones que alteran la secuencia aminoacídica de las enzimas bacterianas. En general, tales mutaciones no afectan a la actividad química de las enzimas. Si lo hicieran, la cepa se extinguiría muy pronto. Ahora bien, mutaciones menores pueden hacer que se incorporen aminoácidos de diferente carga en una enzima, que pueden detectarse después mediante electroforesis. Si colocamos, en un gel, el citoplasma de dos cepas diferentes y aplicamos una corriente eléctrica, las enzimas migrarán a través del gel a velocidades distintas de acuerdo con sus diferentes secuencias aminoacídicas.

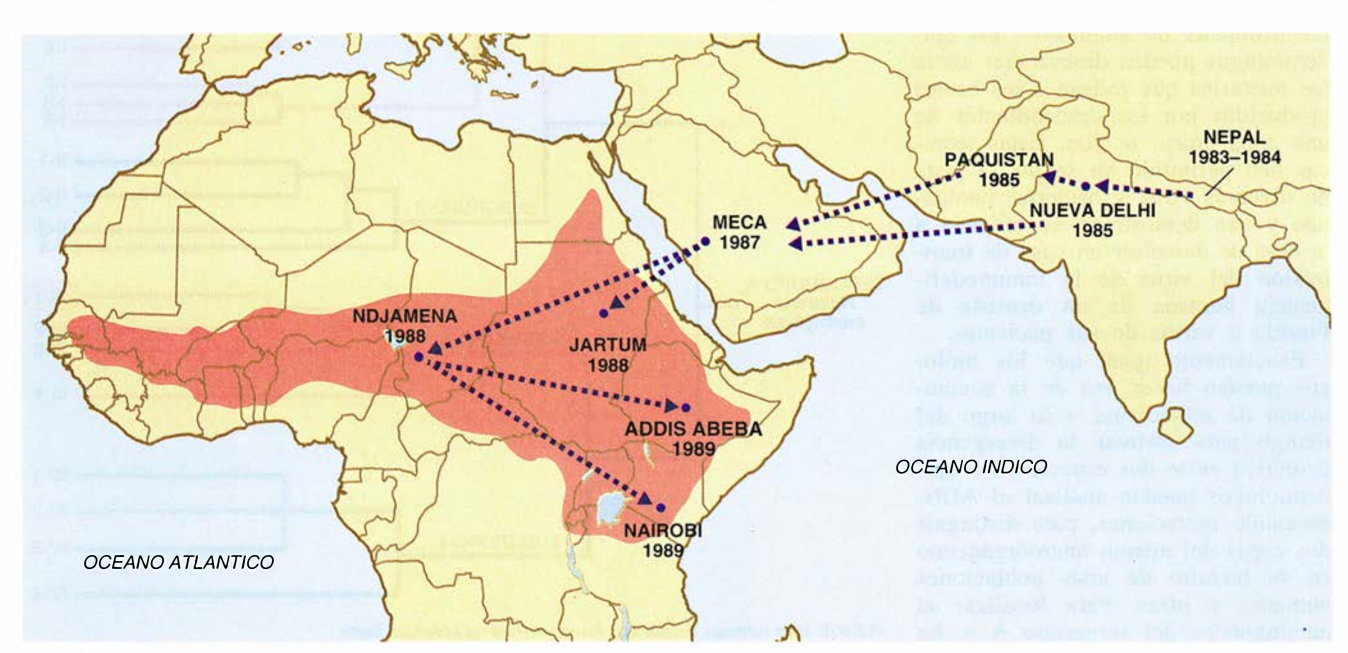
Si dos cepas meningocócicas del serogrupo A se han separado hace poco, la probabilidad de que las mutaciones se hayan acumulado en una enzima determinada es baja. Las enzimas de una y otra cepa presentarán similar comportamiento electroforético. Cuanto más diverjan las cepas, tanto mayor será el número de enzimas con diferente movilidad electroforética entre ellas. Para comparar las enzimas de diversas cepas nos valemos del análisis de grupos, una técnica estadística. Este análisis revela la divergencia genética relativa entre cepas. Cada grupo de cepas semejantes representa un clon único, en el que todos los microorganismos aislados están estrechamente relacionados entre sí y derivan, presumiblemente, de la misma célula ancestral reciente. Los investigadores pueden examinar las cepas y construir el árbol genealógico de los diferentes clones.

En un proyecto ambicioso, Mark Achtman, Tom Olyhoek y Brian A. Crowe, de Berlín, se sirvieron de esa técnica para estudiar 423 cepas de meningococos pertenecientes al serogrupo A. Su análisis permitió esbozar la genética del mismo: dividieron así las cepas en cuatro subgrupos, que a su vez se clasificaron en un total de 21 clones diferentes. Aunque este árbol genealógico se halla lejos de ser completo (muchas cepas de los brotes epidémicos acaecidos en los países en vías de desarrollo no se han conservado y recientemente se ha encontrado un quinto subgrupo) nos proporciona una referencia útil para comparar el potencial epidémico de las cepas.

Es razonable suponer que las cepas meningocócicas diverjan lentamente con el paso del tiempo, a medida que se van propagando entre la población. Los estudios realizados demuestran la coexistencia de cepas diversas en condiciones endémicas. Si todas las cepas tuvieran la misma virulencia y las epidemias se debieran exclusivamente a factores ambientales y dependientes del huésped, las epidemias resultarían ser policlonales, pues todas las cepas gozarían de idéntica probabilidad de inducir la enfermedad. No es así, sin embargo. El grupo de Berlín, así como Musa Hassan-King y Greenwood en Gambia, han descubierto que las epidemias suelen arrancar de un solo clon. Por tanto, la virulencia clonal podría desempeñar también un papel decisivo.

Con la ayuda de Michael Reeves, del CDC, se han estudiado las cepas allí almacenadas procedentes de epidemias declaradas en el sur de Asia, en Africa y en Oriente Medio, en busca de pruebas de virulencia clonal. Se observó una correlación sorprendente. Se empezó por considerar por separado cada epidemia. Del cotejo posterior entre cepas, surgieron conexiones clarísimas. El grupo de Achtman había determinado ya que un clon, el III-1, desencadenó las epidemias que brotaron en China y Nepal a comienzos de los ochenta. Las pautas electroforéticas de las cepas almacenadas en el CDC revelaron que una serie de epidemias manifestadas en China, Nepal, Arabia Saudita y Chad se debían al mismo clon.

El clon III-1 apareció en China en los años sesenta. Una segunda y grave epidemia asoló el valle de Katmandú poco después de restablecer las comunicaciones entre Nepal y el Tíbet en 1984. La misma cepa se propagó por el norte de India y el Paquistán, causando nuevas epidemias. Al parecer permaneció inactiva en el sur de Asia hasta el verano de 1987. Fue entonces cuando el clon III-1 viajó con los peregrinos del sur de Asia hasta La Meca. Los epidemiólogos de Arabia Saudita y del CDC que investigaron la epidemia de La Meca confirmaron que se inició entre esos peregrinos, entre quienes se ob-



6. UNA BACTERIA VIAJERA perteneciente a la cepa III-1 de los meningococos del serogrupo A ha causado hace poco una pandemia internacional. El patrón y el desarrollo cronológico de las epidemias sugieren que los peregrinos transportaron la infección desde Asia —donde se originó— hasta La Meca, y desde allí a los países del cinturón de la meningitis.

servó el índice más elevado de casos. Concluida la peregrinación, los portadores del III-1 regresaron a sus hogares. Fue su retorno a Estados Unidos lo que motivó nuestra investigación en el aeropuerto Kennedy.

Desgraciadamente, muchos peregrinos musulmanes de los países del cinturón de la meningitis también eran portadores del clon III-1. No sorprende, pues, que en 1988 surgiera simultáneamente la epidemia en el Chad y en Sudán. Con el tiempo, la epidemia del III-1 se extendió por el Africa Oriental, afectando a Etiopía, Kenia y Uganda. Los funcionarios de sanidad pública están preocupados ante la posibilidad de que se declaren también epidemias del III-1 en otros países del cinturón de la meningitis. Hace un año hubo una epidemia de meningitis en Togo, pero todavía no se ha demostrado si la originó la cepa III-1.

Aunque esta cepa sea responsable de cientos de miles de casos de meningitis, no parece la única virulenta. Los análisis clonales de las cepas meningocócicas han revelado la intervención de otros clones en epidemias semejantes en Africa y Asia. Lo que induce a pensar que la introducción de un clon potencialmente epidémico en las circunstancias adecuadas puede tener consecuencias devastadoras. Se han propuesto dos explicaciones de este proceso: o bien los clones epidémicos se propagan al azar conforme van asentándose entre la población o bien sobreviven escapando a la inmunidad de rebaño. A semejanza de las epidemias de gripe, se ha sugerido que las de meningitis podrían ser resultado de un cambio antigénico. Aunque todos los meningococos del serogrupo A compartan el mismo polisacárido, cada clon difiere de los demás en otros antígenos expuestos en la superficie del microorganismo. Una vez debilitada la inmunidad frente a los antígenos compartidos, un nuevo clon con antígenos de superficie diferentes escaparía a la vigilancia inmunitaria e iniciaría una epidemia. En este caso, los epidemiólogos que vigilen el desenvolvimiento de la enfermedad verán un "cambio antigénico" a medida que los nuevos clones van sustituyendo a los anteriores.

Si los meningococos experimentan en verdad cambios antigénicos, los ciclos de la enfermedad serían consecuencia, por un lado, del tiempo preciso para que se produzca la pérdida de inmunidad y, por otro, del tiempo medio que tarda un nuevo clon en penetrar en la población. El ambiente también contribuiría en ese caso, ya que la introducción del nuevo clon es insuficiente para iniciar una epidemia. Si el organismo penetra en la población durante la estación lluviosa, puede reforzar la inmunidad. Aunque las condiciones precisas para que se declare una epidemia siguen sin estar claras, parece que, si la cepa penetra en una población cuya inmunidad es baja durante la estación seca, hay un gran riesgo de epidemia.

Así pues, una combinación de factores relacionados con el huésped, el ambiente y el microorganismo parece responsable de la epidemiología peculiar de esta enfermedad. Pero sólo estamos empezando a acotar y ponderar esas notas características. Entra dentro de lo probable que las epidemias no sean uniformes y que haya otros mecanismos que expliquen algunos de los rasgos de los brotes de meningitis, si no todos. Por ejemplo, ¿por qué no produjo la cepa III-1 una epidemia en Estados Unidos? El país no ha padecido epidemias serias desde los años cuarenta, por lo que cabría esperar una inmunidad baja frente a la cepa III-1. Si algo no falta en Estados Unidos son infecciones respiratorias superiores. Tampoco se sabe por qué las infecciones infantiles por *Neisseria* proporcionan una protección tan satisfactoria en los países industrializados y, sin embargo, no ejercen el mismo efecto en el continente africano. Aunque la hipótesis del cambio antigénico es bastante atractiva, se requieren aún estudios a largo plazo en Africa para determinar su validez.

Existe la esperanza de que nuevos descubrimientos ayuden a reducir la amenaza de las epidemias meningocócicas. Las vacunas de que disponemos se basan en la cápsula de polisacáridos que rodea la bacteria y no son eficaces en los lactantes vacunados durante campañas rutinarias. La protección que esas vacunas confieren a los niños es de corta duración; además, su vacunación durante períodos no epidémicos no les protege del siguiente brote.

Podría resolver ese problema una nueva técnica que consiste en la conexión química de los polisacáridos capsulares a un vector proteico, creando una vacuna conjugada. En efecto, se ha mostrado bastante eficaz en los niños una vacuna conjugada de polisacárido y proteína similar frente *Hemophilus influenzae*, otra bacteria productora de meningitis. La Organización Mundial de la Salud ha promovido la investigación para producir y ensayar vacunas conjugadas de este tipo contra el meningococo del grupo A. Otra característica interesante de estas vacunas es su capacidad para reducir el estado de portador del microorganismo entre personas sanas, con lo que se cortaría su transmisión.

De esos y otros nuevos enfoques podrían venir los remedios necesarios para proteger a las poblaciones en riesgo. Pero los obstáculos a los que se enfrenta la creación de nuevas vacunas son de índole política y económica, no menos que científica. El presupuesto sanitario anual para algunos países subdesarrollados no llega a los 5 dólares por habitante, de modo que enfoques creativos para la producción y la adquisición de vacunas son una necesidad urgente.

De momento, lo apremiante es la detección precoz de una epidemia inminente. Aunque la vacuna actual no confiere una inmunidad prolongada, puede usarse en campañas de urgencia durante un brote epidémico. Se están ideando métodos para organizar campañas de vacunación basadas en el número de casos de una determinada población. La detección de nuevos clones podría servir también de sistema de alarma precoz. Jan T. Poolman ha desarrollado anticuerpos monoclonales para la detección de clones meningocócicos del serogrupo A. Ha simplificado las técnicas, lo que permite acelerar el seguimiento de clones.

La difusión de la cepa III-1 es una muestra más del grado de interconexión de la población mundial. Nosotros pudimos eliminar el estado de portador de la cepa III-1 en una escasa fracción de peregrinos estadounidenses que acudieron a La Meca. Por fortuna, los factores socioeconómicos probablemente evitaron que esta cepa diera lugar a brotes en Estados Unidos y en Europa. No hemos sido tan afortunados con otras enfermedades, entre ellas el sida. La imposición de análisis y cuarentenas a los viajeros nunca funcionó bien y está claro que tampoco resultará en el futuro.

BIBLIOGRAFIA COMPLEMENTARIA

THE EPIDEMIOLOGY OF ACUTE BACTERIAL MENINGITIS IN TROPICAL AFRICA. B. M. Greenwood en *Bacterial Meningitis*. Dirigido por J. D. Williams y J. Burnie. Academic Press, 1987.

GLOBAL EPIDEMIOLOGY OF MENINGOCOCCAL DISEASE. Benjamin Schwartz, Patrick S. Moore y Claire V. Broome en *Clinical Microbiology Reviews*, vol. 2, Suplemento, págs. S118-S124; abril de 1989.

MENINGOCOCCAL MENINGITIS IN SUB-SAHARIAN AFRICA: A MODEL FOR THE EPIDEMIC PROCESS. Patrick S. Moore en *Clinical Infectious Diseases*, vol. 14, n.º 2, páginas 515-525; febrero de 1992.

El universo inflacionario autorregenerante

Versiones recientes de la hipótesis inflacionaria describen el universo como un fractal autorregenerante que crea otros universos inflacionarios

Andrei Linde

Si mis colegas y yo andamos en lo cierto, habrá que abandonar la idea de un universo primigenio en bola de fuego, creada en una gran explosión ("big bang"). Proponemos una teoría nueva, basada en una hipótesis de hace quince años y según la cual el universo sufrió una época de inflación cosmológica. Nuestra teoría sostiene que el cosmos creció exponencialmente en una fracción infinitesimal de segundo y que, al final de ese período, continuó su evolución de acuerdo con el modelo de la gran explosión. A medida que se ha ido refinando el modelo inflacionario, se han venido descubriendo consecuencias sorprendentes. Una de ellas constituye un cambio fundamental en nuestra forma de ver el cosmos. Versiones recientes de la teoría inflacionaria afirman que el universo no es una bola de fuego en expansión, sino un inmenso fractal en crecimiento, formado por muchas burbujas en proceso de inflación que producen más burbujas, que a su vez engendran más, y así *ad infinitum.*

Esa forma peculiar de ver el cosmos no nació de una invención caprichosa. Varios investigadores, primero en Rusia y más tarde en Estados Unidos, propusimos la hipótesis inflacionaria en que se basa el nuevo planteamiento para resolver ciertas complicaciones que surgieron de la vieja idea de la gran explosión. En su forma estándar, la teoría de la gran explosión mantiene que el universo nació hace aproximadamente 15.000 millones de años de una singularidad cósmica, un estado en el que la temperatura y la densidad son infinitamente grandes. Por supuesto, no podemos realmente hablar en términos físicos de estas cantidades infinitas. Se supone que las leyes de la física actual no eran aplicables en aquel instante. Estas empezarían a ser válidas a partir del momento en que la densidad del universo descendió por debajo de la densidad de Planck, unos 10^{94} gramos por centímetro cúbico.

Conforme el universo se expandía, se enfriaba progresivamente. Los restos del fuego cósmico primordial todavía nos rodean en forma de la radiación de fondo de microondas, que nos dice que la temperatura del universo ha disminuido hasta unos 2,7 grados Kelvin. El descubrimiento en 1965 de esta radiación de fondo por Arno A. Penzias y Robert W. Wilson, de los laboratorios Bell, apuntaló el afianzamiento de la teoría cosmológica de la gran explosión, modelo que, además, explicaba la cuantía de hidrógeno, helio y otros elementos ligeros en el universo.

Pero conforme se profundizaba en la teoría aparecían los problemas. Por citar un ejemplo: la teoría estándar de la gran explosión, unida a la moderna teoría de las partículas elementales, predice la existencia de muchas partículas superpesadas dotadas de carga magnética, esto es, objetos con un solo polo magnético. Estos polos magnéticos tendrían una masa de 10^{16} veces la del protón, unos 0,00001 miligramos. Según la teoría estándar de la gran explosión, los monopolos habrían surgido en una fase precoz del universo y deberían abundar tanto como los protones. En este caso, la densidad media de materia del universo sería unos 15 órdenes de magnitud mayor que su valor actual, cifrado en unos 10^{-29} gramos por centímetro cúbico.

Este y otros problemas forzaron a los físicos a analizar con más cuidado las hipótesis básicas implícitas en la teoría cosmológica estándar. Y encontramos que muchas eran harto sospechosas. Pasaré revista a seis de las más difíciles. El primer problema, y más importante, es la propia existencia de la gran explosión. ¿Qué hubo antes? Si el espacio-tiempo no existía, ¿cómo pudo todo aparecer de la nada? ¿Qué surgió primero, el universo o las leyes que determinan su evolución? Explicar esta singularidad inicial, dónde y cómo empezó todo, sigue siendo uno de los problemas más arduos de la cosmología moderna.

Otro punto problemático es la planitud del espacio. La teoría de la relatividad general sugiere que el espacio podría ser muy curvo, con un radio del orden de la longitud de Planck, o 10^{-33} centímetros. Vemos, sin embargo, que nuestro universo es prácticamente plano en una escala de 10^{28} centímetros, el radio del universo observable. Este resultado de nuestra observación difiere de las predicciones teóricas en más de 60 órdenes de magnitud.

Una discrepancia similar entre teoría y observación concierne al tamaño de nuestro universo. Las observaciones cosmológicas muestran que nuestra parte del universo contiene al menos 10^{88} partículas elementales. Pero, ¿por qué es tan grande el universo? Si se toma un universo de un tamaño inicial dado por la longitud de Planck, y una densidad inicial igual a la densidad de Planck, se puede calcular, mediante el modelo estándar, cuántas partículas elementales contendría. La respuesta pudiera

ANDREI LINDE es uno de los creadores de la teoría de la inflación cosmológica. Después de licenciarse por la Universidad de Moscú, hizo su doctorado en el Instituto de Física P. N. Lebedev de Moscú, donde empezó a buscar las conexiones entre física de partículas y cosmología. Desde 1990 enseña física en la Universidad de Stanford.

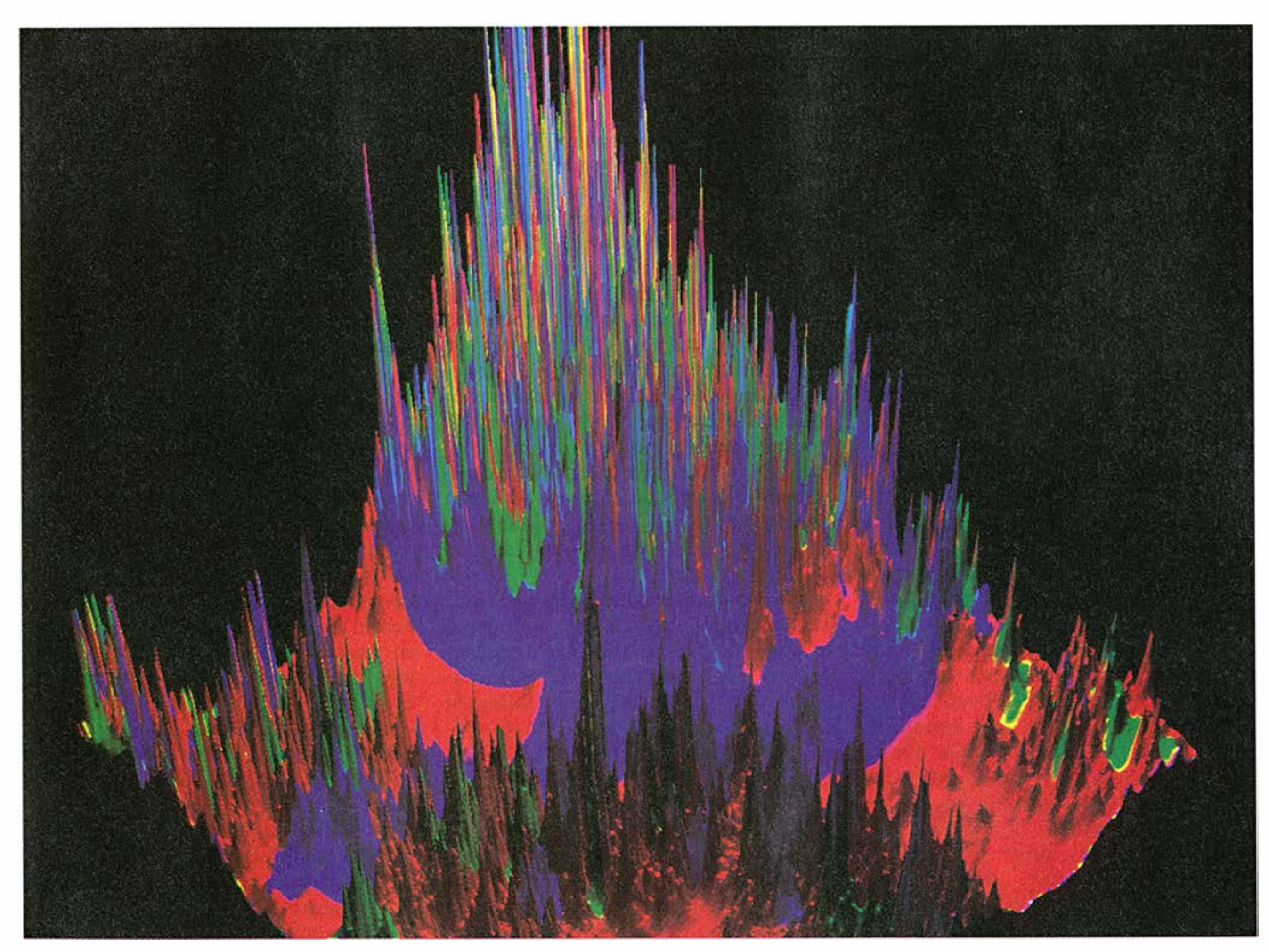
sorprender a más de uno: en todo el universo cabría una sola partícula elemental, o como máximo diez. No podría albergar ni siquiera a un lector de *Investigación y Ciencia*, con sus alrededor de 10^{29} partículas elementales. Algo falla, pues, en la teoría estándar.

El cuarto problema se refiere a la sincronización de la expansión. En el modelo estándar, las diversas partes del universo empezaron a expandirse simultáneamente. Pero, ¿cómo pudieron todas ellas sincronizar el comienzo de su expansión? ¿Quién dio la señal?

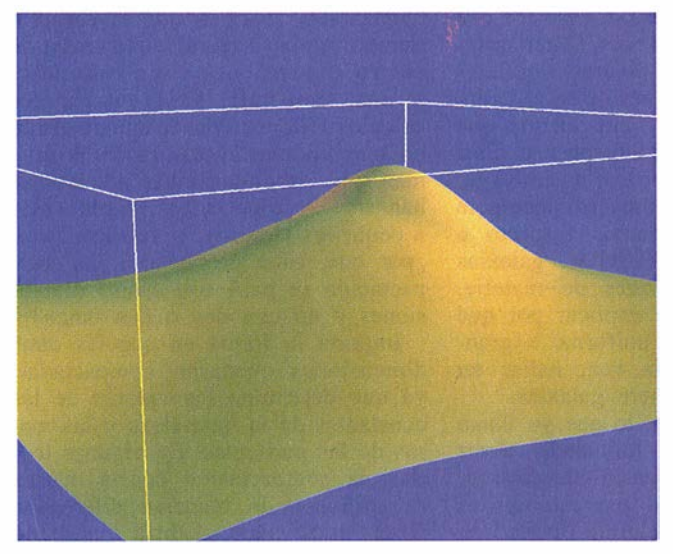
En quinto lugar, está la cuestión de la distribución de materia del universo. A muy gran escala, la materia se extiende con una uniformidad sorprendente. A través de más de 10.000 millones de años-luz, su distribución se aleja de la homogeneidad perfecta en menos de una parte en 10.000. Durante mucho tiempo, nadie tuvo ni idea del porqué de semejante homogeneidad, pero los que no tienen ideas a veces tienen principios. Una de las piedras angulares del modelo estándar reside en el "principio cosmológico", que afirma que el universo debe ser homogéneo. Esta aseveración no sirve, sin embargo, de mucho, pues el universo incorpora desviaciones importantes respecto a la homogeneidad: estrellas, galaxias y otras aglomeraciones de materia. Por tanto, debemos explicar por qué el universo es tan uniforme a grandes escalas y, a la vez, hallar un mecanismo que genere galaxias.

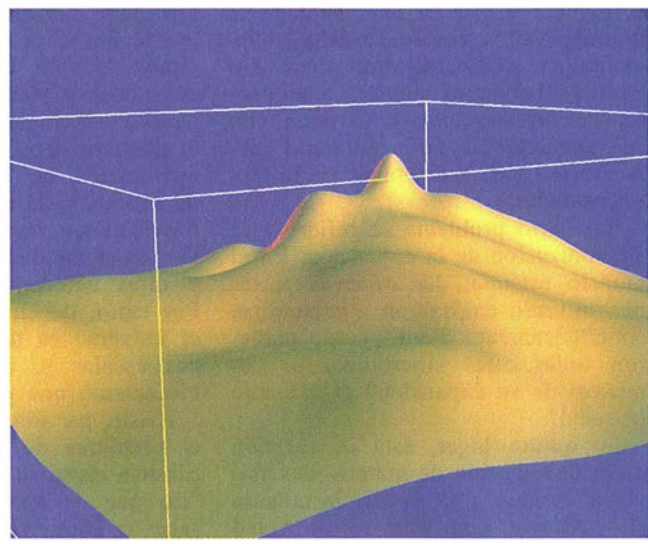
Existe, por último, lo que yo llamo el problema de la unicidad. Albert Einstein captó su esencia cuando dijo: "Lo que de verdad me interesa es saber si Dios tuvo alguna libertad en la creación del mundo." Sin duda, pequeños cambios en las constantes de la naturaleza podrían haber hecho que el universo se desarrollase de forma totalmente distinta. En efecto, muchas teorías actuales de partículas elementales suponen que el espacio-tiempo tenía originalmente más de cuatro dimensiones (tres espaciales y una temporal). Para cuadrar los cálculos teóricos con el mundo físico en que vivimos, estos modelos afirman que las dimensiones adicionales han sido "compactadas", o reducidas a pequeños tamaños, y veladas. Pero, ¿por qué, cabe preguntarse, la compactación se paró con cuatro dimensiones y no con dos o con cinco?

Importa la forma en que las otras dimensiones quedaron compactadas, ya que determina los valores de las constantes de la naturaleza y las masas de las partículas. En algunas teorías, la compactación podría ocurrir de millones de maneras diferentes. Hace unos años, habría resultado absurdo preguntarse por qué el espacio-tiempo tiene cuatro dimensiones, por qué la constante gravitacional es tan pequeña o por qué el protón es casi 2000 veces más pesado que el electrón. Hoy en día, el pro-



1. EL UNIVERSO AUTORREGENERANTE consta, en una simulación de ordenador, de dominios exponencialmente grandes, cada uno de los cuales tiene diferentes leyes de la física (*representadas por colores*). Los picos pronunciados son nuevas "grandes explosiones"; sus alturas corresponden a la densidad de energía del universo en ese punto. En la parte alta de cada pico, los colores fluctúan deprisa, señal de que las leyes de la física no se han asentado allí todavía. Estas se fijan sólo en los valles, uno de los cuales corresponde al tipo de universo donde vivimos.







2. LA EVOLUCION del campo escalar produce muchos dominios inflacionarios, como se observa en esta secuencia de imágenes generadas por ordenador. En la mayor parte del universo, los campos escalares decrecen (*representado en la imagen por va-*

greso de la física de partículas hace que de la respuesta a estas cuestiones dependa nuestro conocimiento de la construcción del mundo.

Todos estos problemas (y otros que no he mencionado) nos dejan perplejos. Por eso resulta alentador ver que muchos admiten solución en el contexto de la teoría del universo inflacionario autorregenerante.

Las características básicas de la hipótesis inflacionaria hunden sus raíces en la física de partículas elementales. Por ello me gustaría llevarles a una breve excursión por este campo, en concreto por la teoría unificada de las interacciones débil y electromagnética. Ambas fuerzas actúan a través del intercambio de partículas. Los fotones median la fuerza electromagnética, mientras que las partículas *W* y *Z* son responsables de la fuerza débil. Pero, si bien los fotones carecen de masa, las partículas *W* y *Z* la tienen grandísima. Para unificar las interacciones débil y electromagnética, los físicos introdujeron los campos escalares.

Los campos escalares no son algo muy usual, pero existe una analogía familiar: el potencial electrostático (el voltaje de un circuito, por ejemplo). Surgen campos eléctricos allí donde este potencial es desigual (como ocurre entre los polos de una pila) o si varía con el tiempo. Si todo el universo tuviera el mismo potencial electrostático, digamos 220 volts, nadie lo notaría; el potencial parecería ser simplemente otro estado de vacío. De forma similar, un campo escalar constante tiene la apariencia del vacío: no lo advertiríamos aunque nos rodeara.

Estos campos escalares llenan el universo y se manifiestan por sus efectos sobre las propiedades de las partículas elementales. Si un campo escalar interacciona con las partículas *W* y *Z*, éstas adquieren masa. Las partículas que no interaccionan con el campo escalar, así los fotones, permanecen ligeras.

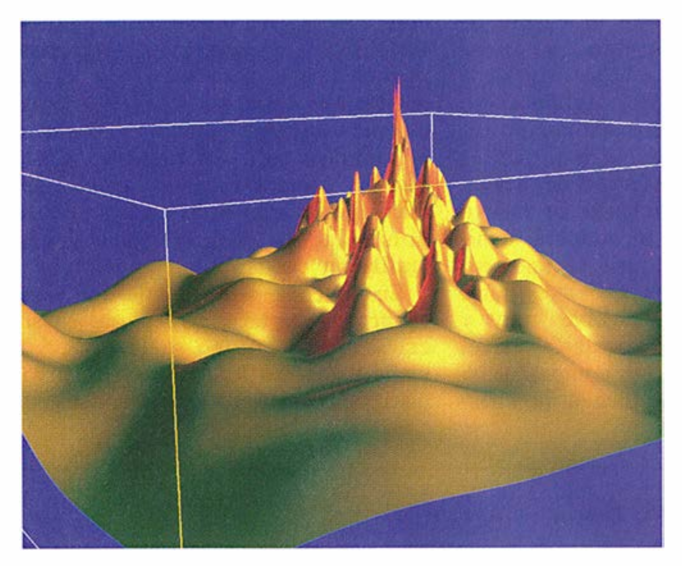
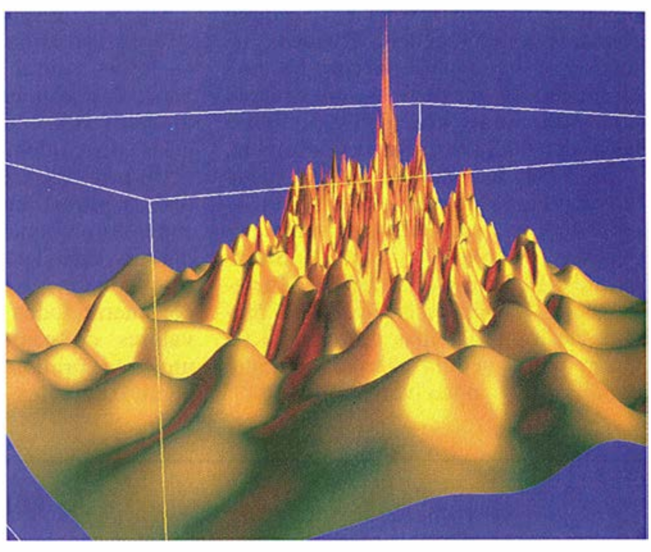
Por tanto, para describir la física de las partículas elementales, se arranca de una teoría en la que inicialmente todas las partículas son ligeras y donde no hay ninguna diferencia fundamental entre las interacciones débil y electromagnética. La diferencia surge sólo más tarde, cuando el universo se expande y se llena de distintos campos escalares. Al proceso por el que las fuerzas fundamentales se diferencian se le conoce por "mecanismo de ruptura de la simetría". El valor particular del campo escalar que aparece en el universo viene determinado por la posición del mínimo de su energía potencial.

Los campos escalares desempeñan un papel fundamental lo mismo en cosmología que en física de partículas, pues proporcionan el mecanismo que genera la rápida inflación del universo. De acuerdo con la teoría de la relatividad general, el universo se expande a un ritmo que viene a ser proporcional a la raíz cuadrada de su densidad. Si el universo estuviera compuesto por materia normal, entonces la densidad disminuiría rápidamente con la expansión cósmica. Por la misma razón, la expansión del universo se frenaría rápidamente conforme disminuyera su densidad. Pero, debido a la equivalencia einsteiniana entre masa y energía, la energía potencial del campo escalar también contribuye a la expansión. En ciertos casos, esta energía decrece mucho más lentamente que la de la materia ordinaria.

La persistencia de dicha energía puede llevar a un estado de expansión celerísima del universo, llamado inflación. Esta posibilidad surge incluso en la versión más simple de la teoría del campo escalar, según la cual la energía potencial alcanza su mínimo allí donde se anula el campo escalar. Por teoría de la gravitación de Einstein, la energía del campo escalar tuvo que haber causado una rápida expansión del universo. Más tarde, la expansión se frenó al llegar el campo escalar al mínimo de su energía potencial.

Una manera de imaginarse la situación es comparándola con una canica rodando por los lados de un cuenco [*figura 3*]. El fondo del cuenco representa el mínimo de energía y la posición de la canica corresponde al valor del campo escalar. Ni que decir tiene que las ecuaciones que describen el movimiento del campo escalar en un universo en expansión son algo más complicadas que las ecuaciones de la canica moviéndose en un cuenco vacío. Aquéllas contienen un término extra correspondiente al rozamiento, o viscosidad. Este rozamiento es similar al que resultaría si hubiera melaza dentro del cuenco. La viscosidad del líquido depende de la energía del campo: cuanto más alta esté la canica en el cuenco, más denso será el líquido. Por tanto, si el campo era muy grande en un principio, la energía disminuyó con lentitud suma.

La lenta disminución de la energía del campo escalar ejerce efectos cruciales en el ritmo de expansión. La

***lles y depresiones*). En otros lugares, las fluctuaciones cuánticas hacen que el campo escalar crezca; en ellos, representados por picos, el universo se expande rápidamente, creándose nuevas regiones inflacionarias. Vivimos en uno de los valles.**

caída fue tan gradual, que la energía potencial del campo escalar se mantuvo casi constante mientras se expandía el universo. Este comportamiento contrasta con el de la materia ordinaria, cuya densidad decrece rápidamente en un universo en expansión. Gracias a la elevada energía del campo escalar, el universo continuó expandiéndose, a velocidades mucho mayores que las predichas por las teorías cosmológicas preinflacionarias. El tamaño del universo durante este régimen creció de manera exponencial.

El estadio de inflación automantenida y exponencialmente rápida no se prolongó demasiado. Duró quizás unos 10^{-35} segundos. Una vez la energía del campo disminuyó, la viscosidad casi desapareció y acabó la inflación. Como la canica cuando llega al fondo del cuenco, el campo escalar comenzó a oscilar alrededor del mínimo de su energía potencial. Conforme oscilaba el campo escalar, perdía energía mediante la creación de partículas elementales, partículas que interaccionaron unas con otras hasta que alcanzaron el equilibrio térmico a cierta temperatura. A partir de ese momento, el modelo estándar de la gran explosión describe de forma correcta la evolución del universo.

La diferencia fundamental entre la teoría inflacionaria y la antigua cosmología se hace patente cuando se calcula el tamaño del universo al final de la inflación. Aun cuando el universo al comienzo de ésta fuera del orden de la longitud de Planck, transcurridos 10^{-35} segundos de inflación, ese dominio adquiere un tamaño increíble. Según algunos modelos inflacionarios, su tamaño en centímetros puede llegar a ser igual a $10^{10^{12}}$, esto es, un uno seguido de un billón de ceros. Aunque estos guarismos dependen del modelo usado, en la mayoría de las versiones este tamaño es muchos órdenes de magnitud mayor que el tamaño del universo observable, 10^{28} centímetros.

Gracias a semejante chorro inflacionario se resuelve de inmediato la mayoría de los problemas del modelo estándar. Nuestro universo parece plano y uniforme porque todas las inhomogeneidades fueron estiradas en un factor $10^{10^{12}}$. La densidad primordial de monopolos y otros "defectos" indeseados se diluyeron exponencialmente. (Hemos descubierto que los monopolos pueden inflarse ellos solos y, por tanto, alejarse por sí mismos de nuestro universo observable.) El universo ha adquirido un tamaño tal, que ahora sólo podemos ver una pequeña fracción del mismo. Por eso, nuestra región del universo parece tan plana, a semejanza de un área pequeña en la superficie de un balón inflado enorme. Ya no es, pues, necesario exigir que todas las partes del universo empiecen a expandirse simultáneamente. Basta un dominio del menor tamaño posible, 10^{-33} centímetros, para producir todo lo que vemos hoy en día.

La teoría de la inflación no nació con semejante sencillez conceptual. El empeño por obtener el estadio de expansión exponencial del universo tiene una larga historia, sólo conocida en parte por los lectores occidentales debido a los avatares de la política.

La primera versión realista de la teoría inflacionaria fue presentada en 1979 por Alexei A. Starobinsky, del Instituto Landau de Física Teórica de Moscú. El modelo de Starobinsky despertó el interés de los astrofísicos rusos, y durante dos años fue el tema principal de discusión en todos los congresos de cosmología de la antigua Unión Soviética. Su modelo, sin embargo, era bastante complicado (se basaba en la teoría de anomalías en la gravedad cuántica) y no aclaraba el comienzo real de la inflación.

En 1981, Alan H. Guth, del Instituto de Tecnología de Massachusetts (MIT), sugirió que el universo caliente, en un estadio intermedio, podría expandirse exponencialmente. Su modelo se basaba en una teoría, propuesta en 1972 por David A. Kirzhnits y por mí en el Instituto P. N. Lebedev de Física de Moscú, que interpretaba el desarrollo del universo primordial como una serie de transiciones de fase. Según esta idea, a medida que el universo se expandía y enfriaba, iba condensándose en diferentes formas. El vapor de agua sufre transiciones de fase similares: al enfriarse, el vapor se condensa en agua, la cual, si sigue enfriándose, se convierte en hielo.

La idea de Guth postulaba que el proceso de inflación ocurría mientras el universo se hallaba en un estado superenfriado inestable. El superenfriamiento es común en las transiciones de fase; por ejemplo, en condiciones adecuadas el agua se mantiene líquida por debajo de cero grados. Por supuesto, el agua superenfriada termina congelándose; este suceso correspondería al final del período inflacionario. El recurso al superenfriamiento para resolver muchos de los problemas de la teoría de la gran explosión revestía enorme interés. Por desgracia, como Guth mismo apuntó, el universo postinflacionario de su

hipótesis modelo resulta extraordinariamente inhomogéneo. Después de investigar su modelo durante un año, él mismo lo rechazó en un artículo publicado junto con Erick J. Weinberg, de la Universidad de Columbia.

En 1982, introduje la que se llamó "nueva hipótesis del universo inflacionario", redescubierta más tarde por Andreas Albrecht y Paul J. Steinhardt, de la Universidad de Pennsylvania. La hipótesis eliminaba muchos de los problemas del modelo de Guth. Pero seguía siendo bastante complicada y poco verosímil.

Un año después me di cuenta de que la inflación es algo que surge de forma natural en muchas teorías de partículas elementales, incluidos los modelos más sencillos del campo escalar discutidos más arriba. No es necesario invocar efectos de gravedad cuántica, transiciones de fase o superenfriamientos, ni siquiera la hipótesis estándar de que el universo originalmente estaba caliente. Basta considerar todos los posibles tipos y valores de campos escalares en el universo primordial y comprobar si alguno de ellos conduce a la inflación. Los lugares donde no se produce inflación se mantienen pequeños, mientras que los dominios donde acontece la inflación terminan siendo exponencialmente grandes y dominan el volumen total del universo. Dado que los campos escalares pueden tomar valores arbitrarios en el universo primordial, yo llamé a esta hipótesis "inflación caótica".

En muchos sentidos, la inflación caótica reviste tal sencillez, que cuesta entender por qué no se le ocurrió a nadie antes. Yo lo atribuyo a motivos psicológicos: los éxitos gloriosos de

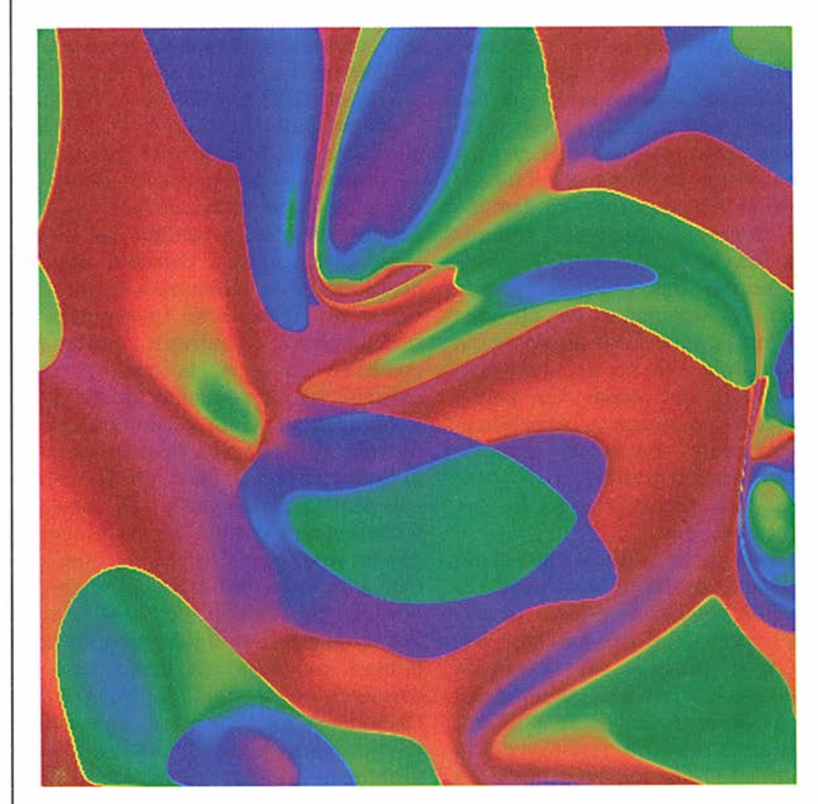
Al octavo día...

La nueva teoría cosmológica, que rompe moldes, se resiste a la representación iconográfica. Una de las principales razones de la popularidad del modelo estándar estriba en la facilidad de imaginarse el universo a la manera de un globo que se expande en todas direcciones. Bastante más difícil es hacerse con la estructura de un universo fractal eternamente autorreproductor. Las simulaciones de ordenador podrían ayudar hasta cierto punto. Aquí describiré algunas de las que he realizado con mi hijo Dmitri.

Empezamos nuestras simulaciones con una sección bidimensional del universo llena de un campo escalar homogéneo. Calculamos la variación del campo escalar en cada punto de nuestro dominio tras el comienzo de la inflación. Más tarde añadimos a este resultado ondas senoidales, correspondientes a fluctuaciones cuánticas que fijan su amplitud.

Iterando el procedimiento, obtuvimos una secuencia de figuras que muestran la distribución del campo escalar en el universo inflacionario. (Por razones de presentación en pantalla, el ordenador contraía la imagen original, en vez de expandir los dominios inflacionarios.) Las imágenes mostraron que en la mayor parte del dominio original el campo escalar decrecía muy despacio [*véase la figura 2*]. Vivimos en esa parte del universo. Pequeñas ondas congeladas encima de un campo casi homogéneo acabarán por dar lugar a las perturbaciones en la temperatura del fondo de radiación descubiertas por el satélite *COBE*. Otras partes de la figura muestran montañas que crecen: éstas corresponden a enormes densidades de energía que originan una rapidísima inflación. Así, se puede interpretar cada pico como una nueva "gran explosión" que crea un nuevo "universo" inflacionario.

La naturaleza fractal del universo resultó incluso más patente al añadir otro campo escalar. Para aumentar el interés del asunto, consideramos una teoría en la que la energía potencial del campo tiene tres diferentes mínimos, representados por distintos colores [*véase la figura 1*]. En una sección bidimensional del universo, los colores cerca de los picos de las montañas cambian sin cesar, indicando que el campo salta rápidamente de uno a otro mínimo. Las leyes de la física allí no están fijas todavía. Pero en los valles, donde el ritmo de expansión es lento, los colores ya no fluctúan. Nosotros vivimos en uno de estos dominios; otros dominios están alejadísimos de nosotros. Las propiedades de las partículas elementales y las leyes de sus interacciones cambian cuando se cruza de uno a otro dominio; uno debería pensárselo dos veces antes de hacerlo.



Un universo de "Kandinsky".

En otro conjunto de figuras exploramos la naturaleza fractal del universo inflacionario siguiendo las ideas de una teoría de partículas distinta. Describir el sentido físico de estas imágenes resulta más difícil. El extraño patrón de color (*izquierda*) corresponde a la distribución de energía en la teoría de axiones (un tipo de campo escalar). Lo llamamos el universo de Kandinsky, en homenaje al famoso pintor abstracto ruso. Visto desde una perspectiva diferente, los resultados de nuestras simulaciones a veces parecen estrellas en explosión.

Realizamos las primeras series de simulaciones hace varios años, tras convencer a la compañía Silicon Graphics, de Los Angeles, para que nos prestara uno de sus más potentes ordenadores por una semana. Preparar las simulaciones fue trabajo duro, y sólo al séptimo día pudimos terminar la serie de cálculos y ver por primera vez todas estas montañas crecientes que representan los dominios inflacionarios. Las sobrevolamos y gozamos de la vista de nuestro

la teoría de la gran explosión hipnotizaron a los cosmólogos. Supusimos que todo el universo fue creado en el mismo instante, que inicialmente estaba caliente y que el campo escalar se encontraba, desde el comienzo, cerca del mínimo de su energía potencial. Una vez empezamos a relajar estas hipótesis, hallamos que la inflación no constituía ningún fenómeno exótico invocado por los teóricos para resolver sus problemas, sino un régimen general que ocurría en una clase muy amplia de teorías de partículas elementales.

Que una dilatación rápida del universo resuelva de una tacada numerosos y graves problemas cosmológicos puede parecer demasiado bonito para ser cierto. En efecto, si todas las inhomogeneidades desaparecieron con la dilatación, ¿cómo surgieron las galaxias? Hay una respuesta sencilla: mientras eliminaba las inhomogeneidades existentes, la inflación creaba otras nuevas.

Estas inhomogeneidades surgen de efectos cuánticos. Según la mecánica cuántica, el espacio vacío no lo está del todo, sino que lo llenan pequeñas fluctuaciones cuánticas; pueden éstas considerarse ondas, u ondulaciones en los campos físicos, que poseen todas las longitudes de onda posibles y se mueven en todas direcciones. No podemos detectar estas ondas directamente; de vida muy breve, son microscópicas.

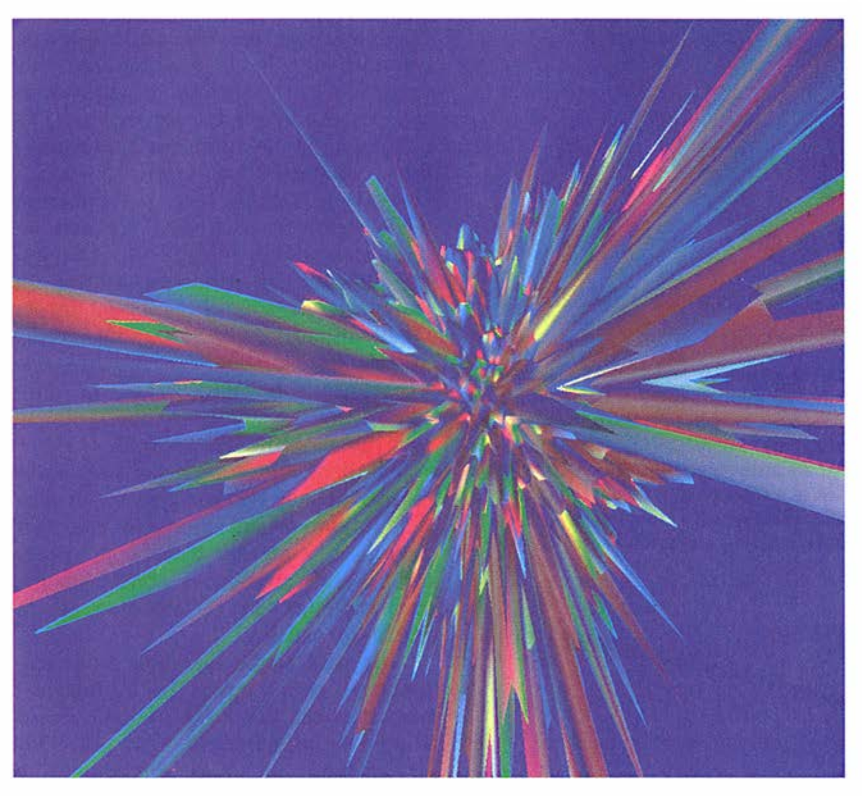
En el universo inflacionario, la estructura del vacío resulta incluso más complicada: la inflación estira rápidamente esas ondas. Cuando sus longitudes de onda alcanzan cierto valor, las ondulaciones empiezan a "sentir" la curvatura del universo. En este momento dejan de crecer por culpa de la viscosidad del campo escalar (recuérdese que las ecuaciones que describen la evolución del campo contienen un término de fricción).

Las primeras fluctuaciones que fijaron su amplitud son las que tenían grandes longitudes de onda. Conforme el universo continuó su expansión, nuevas fluctuaciones se estiraron y congelaron encima de las anteriores. A partir de ese momento, no es lícito llamar fluctuaciones cuánticas a estas ondas, ya que la mayoría tienen longitudes de onda grandísimas. Porque no evolucionan ni desaparecen, estas ondas aumentan el valor del campo escalar en algunas zonas y lo disminuyen en otras, creando así heterogeneidades. Tales perturbaciones del universo son decisivas para la formación posterior de las galaxias.

La teoría inflacionaria no se limita a explicar muchas características de nuestro mundo. Predice cosas importantes. La primera, que el universo debe ser sumamente plano, cualidad que admite comprobación experimental, ya que la densidad de un universo plano guarda relación directa con su velocidad de expansión. Hasta ahora, los datos observacionales son coherentes con esta predicción.

Otra predicción comprobable tiene que ver con las perturbaciones de densidad producidas durante la inflación. Se trata de perturbaciones de la distribución de materia en el universo, que incluso podrían venir acompañadas de ondas gravitacionales. Las perturbaciones de densidad y las ondas gravitacionales dejan su huella en el fondo de radiación de microondas. Provocan que la temperatura de esta radiación difiera algo en puntos del cielo distintos. Esta falta de uniformidad es precisamente lo que encontró el satélite *COBE* ("Explorador de la radiación de fondo cósmica"), descubrimiento que luego corroborarían otros experimentos.

Aunque los resultados del *COBE* están de acuerdo con las prediccio-



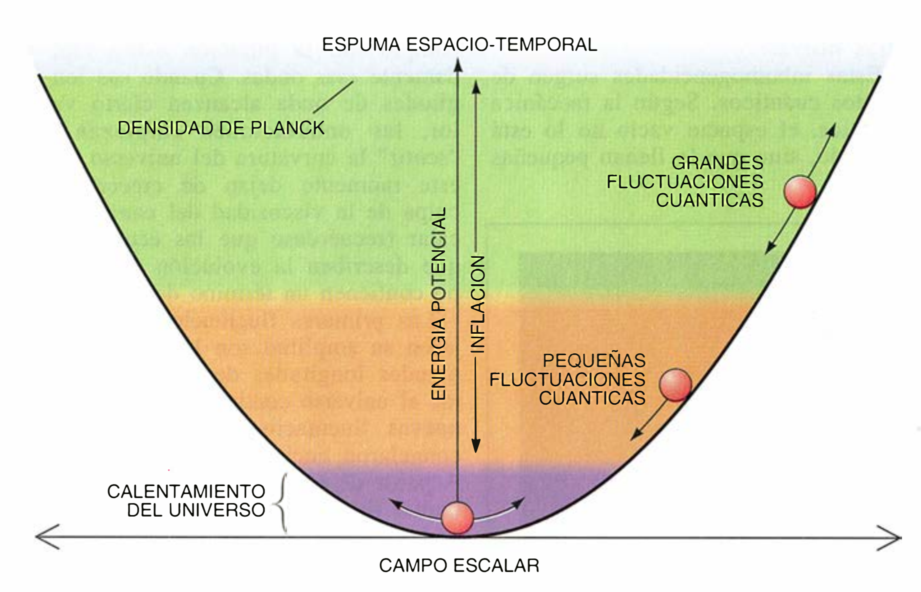
La "explosión" del campo escalar

universo en los primeros momentos de la creación. Mirábamos la brillante pantalla y éramos felices. ¡Vimos que el universo era bello! Pero nuestro trabajo no duró mucho. Al octavo día devolvimos el ordenador y el disco duro de 1 gigabyte cascó, llevándose consigo el universo que habíamos creado. Ahora seguimos con nuestros estudios usando métodos diferentes (y un ordenador Silicon Graphics distinto). Pero se podría jugar a un juego aún más interesante. En lugar de ver el universo en la pantalla de un ordenador, uno podría intentar crear el universo en un laboratorio. Estas ideas son muy conjeturales, pero algunas personas, Alan H. Guth y yo entre ellas, no queremos descartar completamente esta posibilidad. Habría que comprimir algo de materia de tal manera que permitiese que las fluctuaciones cuánticas iniciaran la inflación. Cálculos sencillos en el contexto de la inflación caótica sugieren que menos de un miligramo de materia podría iniciar un universo eternamente autorregenerante.

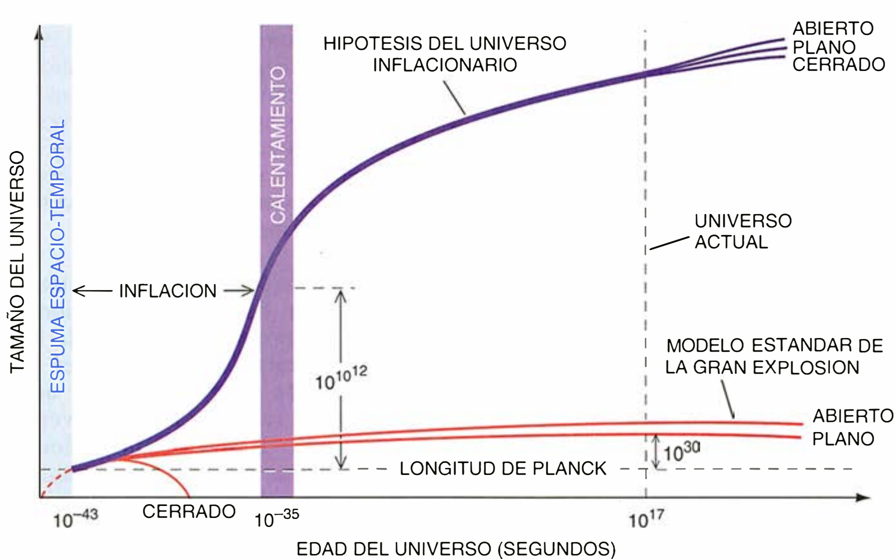
No sabemos si este proceso es posible. La teoría de las fluctuaciones cuánticas que pudieran llevar a un nuevo universo es tremendamente complicada. Y aunque fuera posible "cocinar" nuevos universos, ¿qué haríamos con ellos? ¿Podríamos mandar algún mensaje a sus habitantes, que percibirían su microscópico universo como si fuera tan grande como el nuestro? ¿Es posible que nuestro universo fuera creado por un físico a destajo? Algun día quizás encontremos las respuestas.

nes de la inflación, sería prematuro asegurar que el satélite ha confirmado la teoría de la inflación. Pero sí podrían haber refutado tajantamente los modelos inflacionarios. Y no ha habido tal. No existe, hoy por hoy, ninguna otra teoría capaz de explicar por qué el universo es tan homogéneo y predecir las "arrugas del espacio" descubiertas por el *COBE*.

Pero conviene ser abiertos de mente. Cabe la posibilidad de que nuevos datos observacionales contradigan la cosmología inflacionaria. Por ejemplo, si las observaciones nos dijeran que la densidad del universo



3. REPRESENTACION DEL CAMPO escalar en un universo inflacionario mediante una canica que cae ladera abajo del cuenco. El borde corresponde a la densidad de Planck del universo, por encima de la cual existe una "espuma" espacio-temporal, una región de fuertes fluctuaciones cuánticas. Debajo del borde (*verde*), las fluctuaciones, aunque más débiles, podrían todavía asegurar la autorreproducción del universo. Si la canica se mantiene dentro del cuenco, caerá en la zona menos energética (*naranja*) y se deslizará hacia el fondo lentamente. La inflación termina una vez la canica se ha acercado al mínimo de energía (*violeta*), donde oscila y calienta el universo.



4. EVOLUCION DEL UNIVERSO en el modelo de inflación caótica; difiere de la evolución propuesta por el modelo estándar de la gran explosión. La inflación aumenta el tamaño del universo en un factor $10^{10^{12}}$, de manera que incluso partes de tan sólo 10^{-33} centímetros (la longitud de Planck) llegan a superar el radio del universo observable hoy, 10^{28} centímetros. La inflación predice también que el espacio debe ser esencialmente plano, en el que las líneas paralelas se mantienen paralelas. (Las líneas paralelas se cruzan en un universo cerrado; en un universo abierto, divergen.) Sin embargo, una expansión del tipo gran explosión caliente habría hecho crecer un universo originalmente del tamaño de la longitud de Plank hasta 0,001 centímetros solamente y daría diferentes predicciones sobre la geometría del espacio.

difiere bastante de la densidad crítica, que es la propia de un universo plano, la cosmología inflacionaria se enfrentaría a un auténtico problema. (Sería posible resolverlo si apareciera, pero no sería nada sencillo.)

Hay otra complicación de índole teórica. Los modelos inflacionarios se basan en la teoría de partículas elementales, de la que no podemos decir que esté totalmente asentada. Algunas versiones (en especial, la teoría de supercuerdas) no llevan directamente a la inflación, y obtener la inflación del modelo de supercuerdas podría requerir ideas de nuevo cuño. Deberíamos continuar la búsqueda de teorías cosmológicas alternativas. Pero son muchos los creen que la inflación, o algo similar, es absolutamente esencial para la construcción de una teoría cosmológica coherente. La propia teoría inflacionaria cambia con la evolución de la teoría de partículas. La lista de nuevos modelos incluye la inflación generalizada, la inflación natural, la inflación híbrida y muchos otros. Cada modelo tiene características únicas que pueden contrastarse por experimento u observación. La mayoría, sin embargo, se basan en la idea de la inflación caótica.

Así llegamos a la parte más interesante de nuestra historia, a la teoría del universo inflacionario autorregenerante y eterno. Pese a su carácter general, conduce a consecuencias sorprendentes en el contexto de la inflación caótica.

Dije más arriba que podemos imaginarnos como ondas las fluctuaciones cuánticas del campo escalar en el universo inflacionario. Al principio se mueven en todas direcciones y luego se fijan unas encima de otras. Cada onda incrementa ligeramente el valor del campo escalar en algunas partes y lo disminuye en otras.

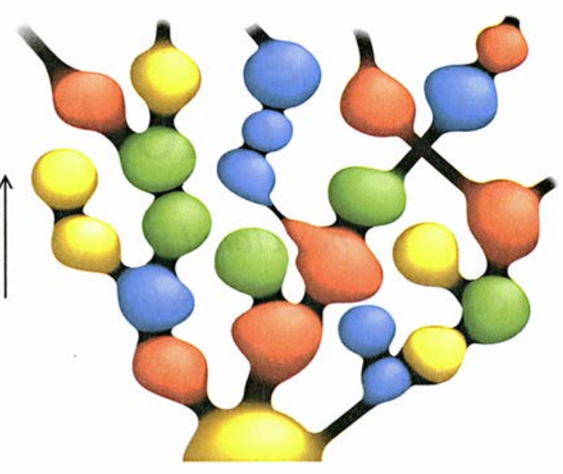
Atendamos ahora a los lugares donde estas nuevas ondas incrementaron de forma persistente el campo escalar. Aunque poco probables, tales regiones existen, y pueden devenir muy importantes. Estos escasos dominios del universo donde el campo salta a un valor bastante alto empiezan a expandirse con velocidad creciente. Cuanto más alto suba el campo escalar, con mayor celeridad se expande el universo; muy pronto, esos dominios poco probables adquieren un volumen mucho mayor que el resto.

De esta teoría se infiere que, si el universo contenía al menos un dominio inflacionario de tamaño suficiente, producirá sin cesar nuevos dominios inflacionarios. La inflación en

cada punto particular puede acabar muy deprisa, pero muchos otros puntos continuarán su expansión. El volumen total de todos estos dominios crecerá sin fin. En esencia, un universo inflacionario crea otras burbujas inflacionarias, que a su vez producen más burbujas.

Llamé inflación eterna a este proceso. Se mantiene como una reacción en cadena y produce una estructura de tipo fractal. En esta hipótesis, el universo, entendido en su globalidad, es inmortal; cada región particular del mismo podría haber surgido de una singularidad en algún momento del pasado y acabar en otra singularidad en el futuro. Sin embargo, no hay fin para la evolución del universo en cuanto tal.

Menos clara aparece la cuestión del origen. Existe la posibilidad de que todas las partes del universo fueran creadas simultáneamente en una singularidad inicial de tipo de la gran explosión, aunque la necesidad de esta hipótesis deja de ser obvia. Más aún, el número total de burbujas inflacionarias en nuestro "árbol cósmico" crece exponencialmente con el tiempo. Por tanto, la mayoría de las burbujas (incluida nuestra parte del universo) crecen ilimitadamente alejadas del tronco de ese árbol. Si bien este modelo hace que la existencia de la gran explosión inicial sea casi irrelevante a todos los efectos prácticos, uno puede considerar el momento de formación de cada burbuja inflacionaria como una nueva "gran explosión". Desde esta perspectiva, la inflación no forma parte del modelo estándar, como creíamos quince años atrás; antes bien, la gran explosión es parte del modelo inflacionario.



5. EL COSMOS AUTORREGENERANTE aparece como un extenso enramado de burbujas inflacionarias. Los cambios de color representan "mutaciones" en las leyes de la física respecto a las de los universos progenitores. Las propiedades del espacio en cada burbuja son independientes del momento en que la burbuja se formó. En este sentido, el universo, entendido en su globalidad, podría ser estacionario, aunque el interior de cada burbuja sea descrito por la teoría de la gran explosión.

Al considerar el proceso de autoregeneración del universo, no podemos evitar recurrir a las analogías, por muy superficiales que parezcan. Y preguntarnos si este proceso no evoca nuestra propia historia. Nacimos hace tantos años. Moriremos. Desaparecerá el mundo complejo de nuestras ideas, sentimientos y recuerdos. Pero también existieron quienes nos precedieron, y habrá quien nos suceda. La propia humanidad, si es lo bastante lista, podrá vivir por mucho tiempo. La teoría de la inflación sugiere un proceso similar en el universo. Uno siente cierto optimismo al saber que, aunque nuestra civilización muera, siempre habrá otros lugares en el cosmos donde la vida surgirá una y otra vez, en todas las formas posibles.

¿Podrían ser las cosas todavía más interesantes? Podrían. Hasta ahora, hemos considerado el modelo inflacionario más sencillo, con un solo campo escalar que tiene un solo mínimo de su energía potencial. Pero los modelos de partículas elementales proponen muchos tipos distintos de campos escalares. Por ejemplo, en las teorías unificadas de las interacciones débil, fuerte y electromagnética existen al menos otros dos campos escalares. La energía potencial de estos campos escalares puede presentar varios mínimos distintos. Ello quiere decir que la misma teoría puede tener distintos "estados de vacío", correspondientes a diferentes tipos de ruptura de simetría entre las interacciones fundamentales y, en consecuencia, a diferentes leyes de baja energía. (Las interacciones a muy alta energía son insensibles al mecanismo de ruptura de la simetría.)

Estas complejidades del campo escalar sugieren que, después de la inflación, el universo podría quedar dividido en dominios exponencialmente grandes, cada uno con leyes físicas de baja energía diferentes. Nótese que tal división ocurre aunque al principio el universo comience en el mismo estado, correspondiente a uno de los posibles mínimos de la energía potencial. En efecto, las fluctuaciones cuánticas grandes pueden hacer que los campos escalares salten fuera del mínimo; es decir, sacar del cuenco una de las canicas y lanzarla a otros cuencos, donde cada cuenco está asociado a diferentes leyes de las interacciones entre partículas. En algunos modelos inflacionarios, las fluctuaciones cuánticas son tan fuertes, que podría incluso variar el número de dimensiones de espacio y tiempo.

Si este modelo es correcto, entonces la física no nos dará, por sí sola, una explicación completa de todas las propiedades de la región del universo que nos ha tocado habitar. La misma teoría física podría dar lugar a regiones enormes del universo que poseyesen propiedades muy distintas. Según este modelo, nos hallamos dentro de un dominio tetradimensional que tiene nuestro tipo de leyes físicas, no porque dominios con distinta dimensionalidad y propiedades distintas sean imposibles o improbables, sino por la sencilla razón de que nuestro tipo de vida no podría existir en otros dominios.

¿Quiere esto decir que entender todas las propiedades de nuestra región del universo requerirá, aparte de conocimientos de física, una investigación profunda de nuestra propia naturaleza, quizás hasta de nuestra consciencia? Esta conclusión sería sin duda una de las más inesperadas que se puedan sacar de los desarrollos recientes de la cosmología inflacionaria.

La evolución de la teoría inflacionaria ha engendrado un paradigma cosmológico nuevo. En la nueva hipótesis, el universo parece ser a la vez caótico y homogéneo, en expansión y estacionario. Nuestra casa cósmica crece, fluctúa y se reproduce eternamente de todas las maneras posibles, como si se ajustara por sí misma a todas las posibles formas de vida que en ella puedan habitar.

Algunas partes de esta nueva teoría, esperamos, permanecerán con nosotros por muchos años. Muchas otras habrán de modificarse para adaptarse a nuevos datos observacionales y a la siempre cambiante teoría de las partículas elementales.

BIBLIOGRAFIA COMPLEMENTARIA

INFLATION AND QUANTUM COSMOLOGY. Andrei Linde. Academic Press, 1990.

PARTICLE PHYSICS AND INFLATIONARY COSMOLOGY. Andrei Linde. Harwood Academic Publishers, 1990.

FROM THE BIG BANG THEORY TO THE THEORY OF A STATIONARY UNIVERSE. A. Linde, D. Linde y A. Mezhlumian en *Physical Review D*, vol. 49, n.º 4, págs. 1783-1826; febrero 1994.

Claves genéticas del desarrollo floral

Los genes que portan información posicional indican a las células de las flores en qué órganos deben convertirse. Un modelo basado en media docena de tales genes prevé el modo en que las mutaciones afectarán a la estructura floral

Elliot M. Meyerowitz

Las flores se componen de varios órganos distintos. Pueden tener sépalos (estructuras foliáceas modificadas), pétalos, estambres (productores de polen) y carpelos (subunidades de los ovarios de una planta). En casi todas las flores silvestres que disponen de una dotación completa de tales órganos, su orden desde la periferia hacia el centro de la flor es sépalos-pétalos-estambres-carpelos. Pero cada especie y cada variedad se distingue por su morfología peculiar y por la disposición en que se distribuyen tales componentes.

El orden de los órganos florales en la planta es rasgo que se hereda. Lo que significa que la información que lo señala debe portarse en su material genético. Pero, ¿cómo se puede codificar la estructura tridimensional de una flor —o de cualquier otro organismo, puestos a ello— en el orden unidimensional de los pares de bases del ADN cromosómico?

Aunque conocemos sólo en parte los mecanismos de la creación de formas durante el desarrollo, sí está claro que la posición relativa que ocupa una célula en el organismo indica a ésta y a su progenie el tipo de estructura que debe formar. Esta información posicional viene vehiculada a través de combinaciones o gradientes de proteínas reguladoras que hay en regiones específicas del organismo.

Atraído por su simplicidad, me centré en la génesis de la forma floral. Las flores constan de cuatro clases de órganos fácilmente distinguibles; cada uno se compone de un pequeño número de tipos de células, y las plantas pueden mantenerse y manipularse sin dificultad en el laboratorio. Partimos de un planteamiento genético: alteramos el ADN de la planta y observamos los efectos en la estructura final de las flores; después, extraemos hipótesis sobre los mecanismos que actúan, ideas que sometemos luego a comprobación experimental en ensayos, genéticos y bioquímicos, que pueden incluir el diseño y producción de nuevos tipos de flores.

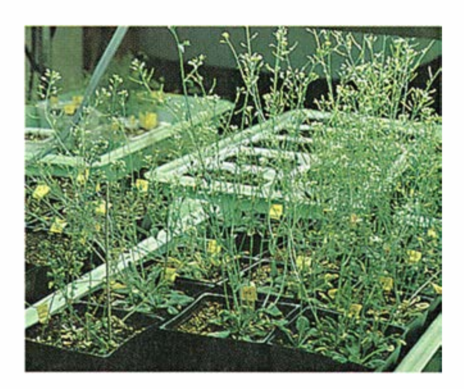
La capacidad de reordenar y transformar a voluntad las flores no puede considerarse asunto menor en un mundo donde la mayor parte de la alimentación humana se funda en productos relacionados con las flores: semillas o frutos. Cuando llegue el día en que dominemos el desarrollo de la forma floral, podremos crear plantas con redoblado rendimiento en estructuras que nos sean más preciadas. Nuestro laboratorio ha dado ya algunos pasos en esa dirección.

Antes de comenzar los experimentos sobre generación de formas florales, sabíamos por adelantado que el enfoque genético tendría éxito. Los expertos en mejora vegetal venían alterando genéticamente muchas variedades de plantas para que ofrecieran un número o una disposición anormal de sus órganos. En particular, las plantas ornamentales más cotizadas son resultados monstruosos de mutaciones genéticas que han alterado el proceso normal de generación de forma: las rosas silvestres sólo tienen cinco pétalos, mientras que la rosa de té híbrida de mi jardín cuenta de 35 a 40; las camelias silvestres presentan estambres y carpelos (como debe ser, para producir semillas), pero yo cultivo una variedad sin unos ni otros y con pétalos ocupando la plaza reservada a dichos órganos reproductores.

De tales variantes genéticas se desprende que las alteraciones de la secuencia del ADN de las plantas modifican persistentemente el orden normal de los órganos. Podríamos, pues, aprender mucho sobre genética del desarrollo de la forma floral investigando variedades mutantes. Pero ni rosas ni camelias nos sirven para un análisis genético riguroso, que precisa el examen de muchísimos individuos. El largo tiempo de generación y el tamaño de estas plantas las hacen prohibitivas para un laboratorio de investigación básica.

Empezamos por elegir el sujeto experimental. Por su tamaño pequeño y rápido crecimiento escogimos *Arabidopsis thaliana* o berro de oreja de

ELLIOT M. MEYEROWITZ, profesor de biología en el Instituto de Tecnología de California, estudia los mecanismos celulares y moleculares de formación de órganos en el desarrollo floral. Inició sus estudios en la Universidad de Columbia. Dedicó su tesis doctoral, en Yale, y su trabajo posdoctoral, en Stanford, a la biología del desarrollo de *Drosophila melanogaster*, la mosca del vinagre. Sin embargo, tras su incorporación en el Caltech, en 1980, quedó fascinado por el desarrollo de las plantas y reorientó sus investigaciones hacia *Arabidopsis*.

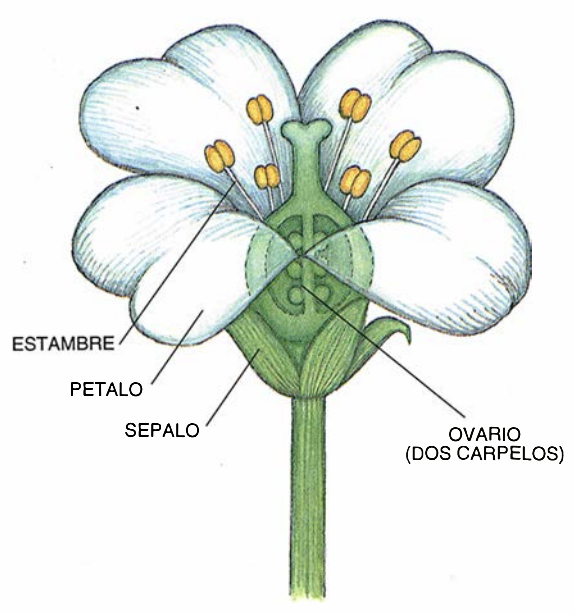


1. FLORES de *Arabidopsis thaliana*, o berro de oreja de ratón (*a la derecha, página siguiente*). Son un material excelente para el estudio genético del desarrollo de los órganos florales. Merced a su exiguo tamaño, pueden cultivarse miles de ellas en una poyata de laboratorio.



ratón. Esta planta de la familia de la mostaza crece bien en interiores bajo la luz fluorescente normal; en una poyata de laboratorio podemos criar varios miles hasta la madurez. Abunda en prados y jardines, con racimos de florecitas blancas que maduran en frutos alargados llenos de semillas diminutas. A pesar de su tamaño, *Arabidopsis* es, en todo lo fundamental, una típica planta con flores.

Las flores normales de *Arabidopsis* tienen cuatro regiones concéntricas, o verticilos, ocupado cada uno por un tipo diferente de órgano floral. En la periferia hay cuatro sépalos verdes, que rodean un segundo verticilo de cuatro pétalos blancos. Dentro de los pétalos están el tercer verticilo, con seis estambres (cuatro largos y dos cortos), y el verticilo más interno, que es un ovario compuesto por dos carpelos soldados.



2. ORGANOS DE LA FLOR. Son de cuatro tipos básicos: sépalos foliáceos, pétalos, estambres (órganos reproductores masculinos) y carpelos (órganos reproductores femeninos). En muchas variedades de flores, estos órganos se disponen en cuatro verticilos concéntricos.

La flor arranca su desarrollo desde una fase de primordio, o proliferación de células indiferenciadas, cerca del ápice de un tallo florífero. A través de la combinación de división celular, diferenciación celular y elongación celular —procesos que ocurren en posiciones precisas—, el primordio inicial se convierte en flor. ¿Cómo se enteran las células primordiales del lugar al que están destinadas?

En los ensayos iniciales, buscamos razas mutantes que presentaran anormal el orden de órganos. Confiábamos en identificar genes individuales cuyos productos proteínicos fueran esenciales para la generación de formas: sin ellos, las células indiferenciadas de una flor en desarrollo malinterpretarían su posición de destino y se transformarían en órganos normales en posición inadecuada. Para conseguir muchas razas de *Arabidopsis*, tratamos las semillas con sulfonato de etilmetano (un mutágeno) y pedimos variedades con flores anormales a Maarten Koornneef, de la Universidad de Agricultura de Wageningen, y otros colegas.

En un comienzo, todas las flores mutantes observadas encajaban en tres grupos. El primero, producido por las mutaciones que inactivan el gen *APETALA2*, tiene carpelos en vez de sépalos en el primer verticilo y estambres en lugar de pétalos en el segundo (por convención, los nombres de los genes de *Arabidopsis* se escriben con letras mayúsculas; las plantas que tienen mutaciones en esos genes llevan los mismos nombres, pero en minúscula).

Un segundo grupo está causado por mutaciones en el gen *APETALA3* o en el *PISTILLATA*. Estas flores presentan sépalos en el primer y segundo verticilos y carpelos en el tercero y cuarto.

En el tercer grupo de mutantes se bloquea el gen *AGAMOUS* [ágamo]. De ello resulta otro orden anormal de órganos, el de la secuencia sépalos-pétalos-pétalos-sépalos. Estos mutantes *agamous* también tienen verticilos internos adicionales de órganos que repiten la disposición de los verticilos externos. La presencia de estas piezas supernumerarias sugiere que el producto del gen *AGAMOUS*, además de establecer la identidad de los órganos de los verticilos tercero y cuarto, interrumpe el crecimiento floral una vez desarrollado el cuarto verticilo.

Cada clase de mutantes reseñada carece de una actividad genética, diferente e imprescindible para un desarrollo normal de la forma floral. Llamaré A a la actividad que falta en los mutantes *apetala2*, B a la actividad ausente en los mutantes *apetala3* y *pistillata*, y C a la actividad ausente en los mutantes *agamous*.

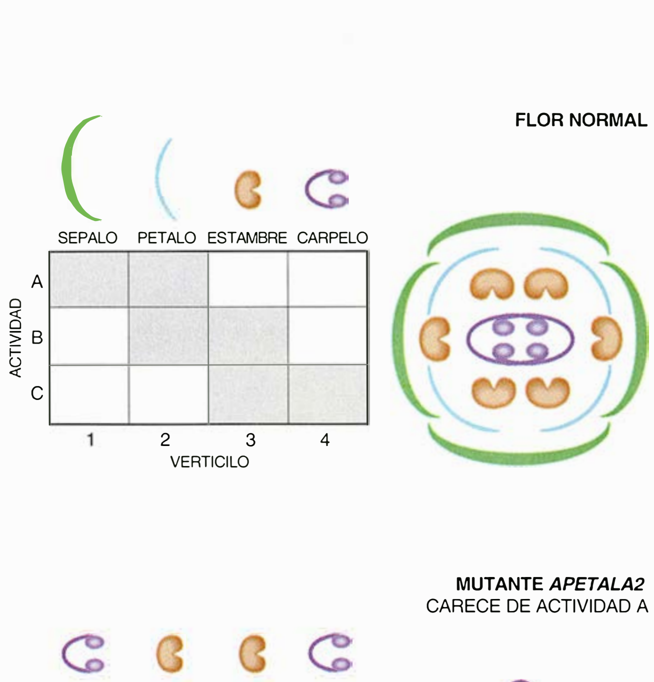
Pese a tan parca información, logramos trenzar una sencilla hipótesis provisional, o modelo de trabajo, sobre la posible contribución normal de esos cuatro genes a la especificación de la identidad de órganos en las flores en desarrollo. El modelo consta de tres partes. Primera, la proteína codificada por cada uno de estos genes está activa en el comienzo de la región afectada por mutaciones en sus genes. Por consiguiente, el gen *APETALA2* de la actividad A actúa en los verticilos primero y segundo, los genes *APETALA3* y *PISTILLATA* de la actividad B actúan en el segundo y tercer verticilos, y el gen *AGAMOUS* de la actividad C actúa en el tercer y cuarto verticilos.

Segunda, las combinaciones de estos productos de los genes determinan la identidad de los órganos florales que se desarrollarán a partir del primordio. Cualquier órgano que se constituya en una región donde sólo esté presente la actividad A se convertirá en un sépalo. Cualquier órgano que se forme donde estén presentes a la vez la actividad A y la B será un pétalo. La presencia combinada de las actividades B y C crea estambres. Si sólo está presente la actividad C, el órgano que emergerá será un carpelo.

La tercera parte del modelo establece la exclusión mutua de las actividades A y C: una región donde esté presente la actividad A carecerá de actividad C, y viceversa. Si se elimina la actividad A (como en los mutantes *apetala2*), C intervendrá, anormalmente, en los verticilos primero y segundo. Pero si falta C (como en los mutantes *agamous*), A estará anormalmente presente en los verticilos tercero y cuarto.

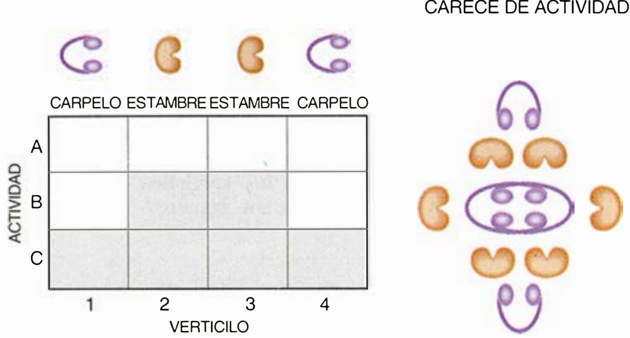
Bastan estas tres reglas para explicar las flores producidas por los mu-

3. DE LA INVESTIGACION GENETICA de flores procedentes de variedades normales y mutantes de *Arabidopsis* ha emergido un modelo simple que explica la disposición de los órganos. Basta la combinación de las tres actividades genéticas —A, B y C— para especificar la identidad de los órganos en cada verticilo. Se ilustran diagramas de las flores y gráficos de las actividades genéticas de cada verticilo en algunas variedades de *Arabidopsis*. En las flores normales, A está activo en los verticilos 1 y 2, B en los verticilos 2 y 3, y C en los verticilos 3 y 4. Las actividades A y C son complementarias: si falta una, la otra se manifiesta en todos los verticilos. Los sépalos sólo crecen en aquellos tejidos en los que únicamente esté presente A; los pétalos resultan de una combinación de A y B. Los estambres se desarrollan donde coinciden B y C, y se producen carpelos donde sólo C está activa.



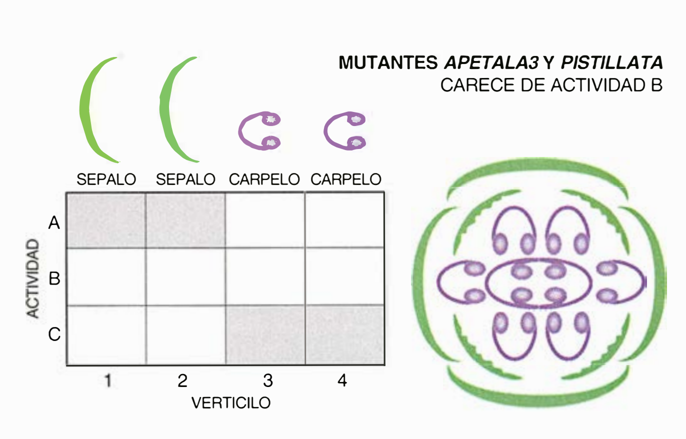
FLOR NORMAL





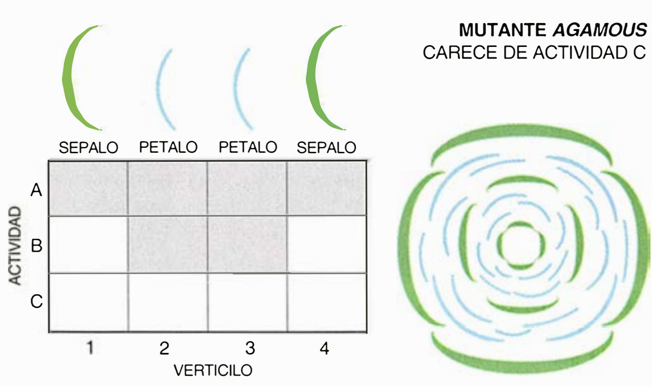
MUTANTE *APETALA2*
CARECE DE ACTIVIDAD A





MUTANTES *APETALA3* Y *PISTILLATA*
CARECE DE ACTIVIDAD B

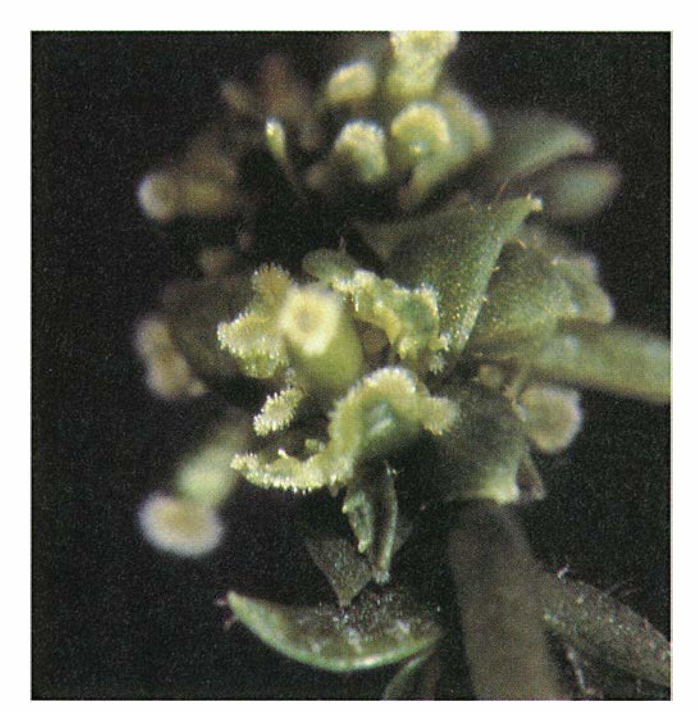


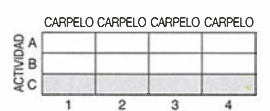




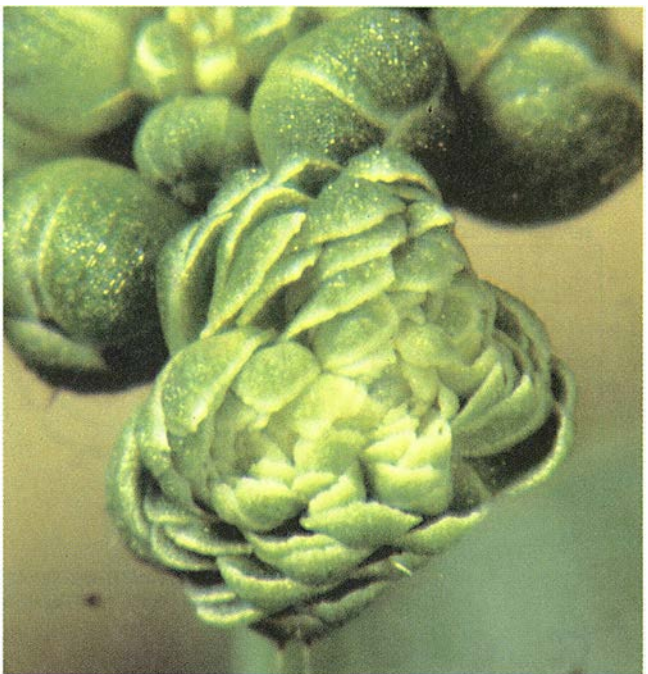
MUTANTE *AGAMOUS*
CARECE DE ACTIVIDAD C







4. LOS DOBLES MUTANTES *apetala2 apetala3* carecen de las dos actividades A y B. La actividad C se extiende a todos los verticilos, y todos los órganos son carpelos.



SEPALO SEPALO SEPALO SEPALO
ACTIVIDAD
A
B
C
1 2 3 4

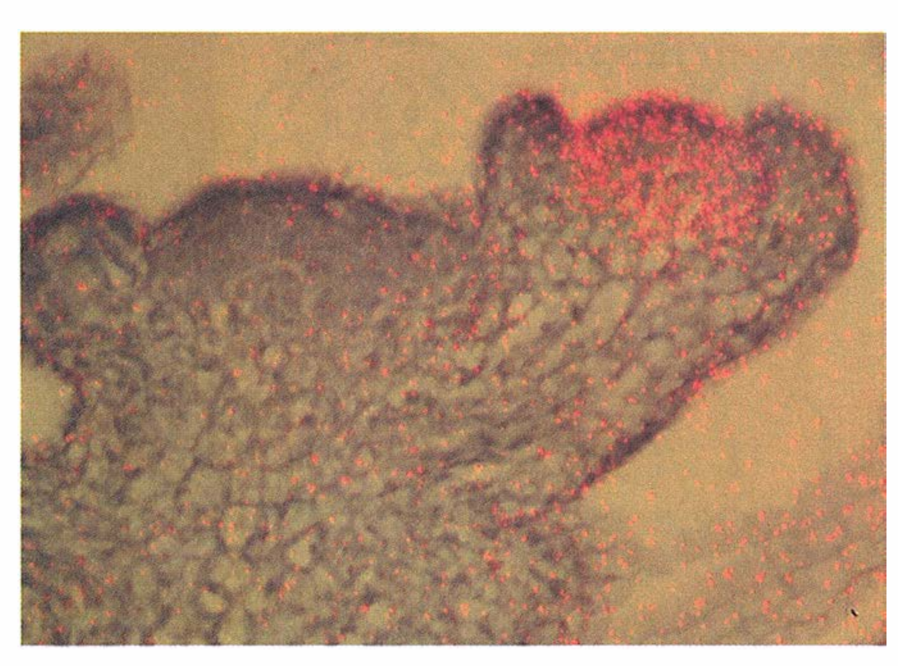
5. LOS DOBLES MUTANTES *apetala3 agamous* carecen de las dos actividades B y C. La actividad A se extiende a todos los verticilos, y así todos los órganos son sépalos.

tantes. Ahora bien, el que un modelo coincida con ciertos datos observados no constituye una demostración de que sea correcto; el modelo debe someterse a contrastación.

El banco de pruebas nos lo proporcionan unos experimentos genéticos muy sencillos. Prevé el modelo qué sucederá si eliminamos pares de actividades. Si se suprimen las actividades A y B —seleccionando plantas con mutaciones en los dos genes *APETALA2* y *APETALA3*, por ejemplo—, el primordio floral sólo tendrá actividad C (que se extenderá a los verticilos primero y segundo, pues falta la actividad A). Por tanto, el modelo predice que todos los órganos florales serán carpelos. Cuando llevamos a cabo el experimento, se confirmó la predicción. Por un razonamiento parecido, si eliminamos a la vez B y C, como en la combinación doblemente mutante *apetala3 agamous*, todos los órganos florales deberían constituirse en sépalos. Y lo hacen.

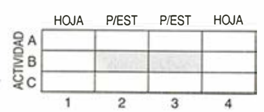
Contando sólo con los tipos de flores observados, no podíamos saber por adelantado qué tipo de órganos aparecerían en plantas mutantes que sólo tuvieran la actividad B. Pero el modelo todavía era capaz de predecir que los órganos de los verticilos tercero y cuarto deberían ser del mismo tipo y, además, diferentes de los órganos del segundo y tercer verticilos. Cuando obtuvimos mutantes dobles *apetala2 agamous* (que carecen de las actividades A y C), descubrimos que tenían hojas en los verticilos primero y cuarto; y órganos intermedios entre pétalos y estambres en los verticilos dos y tres.

Más aún, en flores de mutantes triples sin actividad en ninguna de las tres clases funcionales de genes, todos los órganos florales se resolvieron en hojas. Estos resultados consolidan el modelo e indican que las actividades A, B y C son suficientes para especificar no sólo las diferencias entre los órganos florales, sino

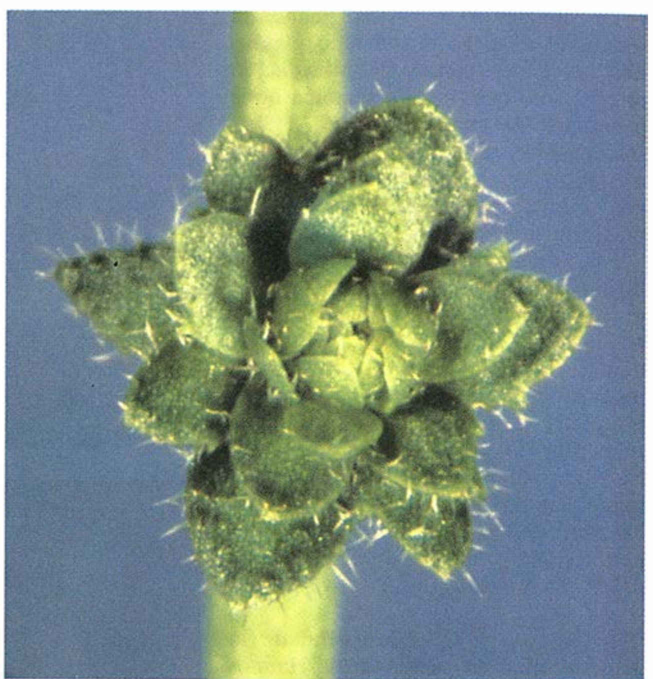


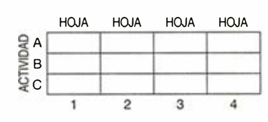
6. EXPERIMENTOS DE HIBRIDACION *in situ* confirman las predicciones del modelo genético al mostrar dónde están activos los genes de identidad de órganos. Las moléculas radiactivas de la sonda se unen al ARN producido por genes específicos que se hallan activos. A la izquierda, las sondas para el ARN *AGAMOUS* (*rojo*) se





7. LOS DOBLES MUTANTES *apetala2 agamous* tienen hojas en los verticilos 1 y 4 y órganos intermedios entre pétalos y estambres en los verticilos 2 y 3.





8. LOS MUTANTES TRIPLES que carecen de las tres categorías de actividad genética tienen flores que consisten enteramente en hojas dispuestas en verticilos.

también entre los órganos florales y los órganos vegetativos: las hojas.

Conseguimos más pruebas de la validez de nuestro modelo de forma radial cuando clonamos algunos de los genes que especifican la identidad de los órganos de la flor. La disponibilidad de esos clones nos permitió llevar a cabo experimentos basados en la técnica de hibridación *in situ*. Con los clones por molde, creamos moléculas de ARN "antisentido", complementarias del ARN producido por los genes. Estas moléculas antisentido contenían pequeñas cantidades de elementos radiactivos.

Cuando se aplicaron a secciones de tejido floral, las sondas antisentido se enlazaban en cualquier ARN complementario procedente de genes activos celulares. Al aplicar una emulsión fotográfica sobre el tejido exponiéndola a la radiación de las sondas, descubrimos dónde se unían estas sondas y, por tanto, dónde estaban activos los genes de interés.

El primer gen que pudimos clonar fue el de la actividad C, *AGAMOUS*. La hibridación *in situ* de sondas *AGAMOUS* en secciones de flores en desarrollo verificó una predicción del modelo: el ARN de *AGAMOUS* aparece muy pronto en el desarrollo de la flor y lo hace sólo en aquellas regiones del primordio que más adelante se convertirán en los verticilos tercero y cuarto.

De acuerdo con nuestro modelo, si falta la actividad A, la actividad C debe extenderse anormalmente a los verticilos primero y segundo. De nuevo, la hibridación *in situ* confirmó algo que preveía el modelo. Cuando sometimos a prueba las flores mutantes *apetala2*, encontramos ARN del tipo *AGAMOUS* en los cuatro verticilos de las flores en desarrollo. Más adelante, al clonar los genes *APETALA3* y *PISTILLATA*, preparamos sondas similares para la actividad B. Según habíamos anticipado, estos genes se hallaban activos en los verticilos segundo y tercero.

localizan dentro del primordio de una flor normal en las células que más tarde formarán parte de los estambres y los carpelos. A la derecha, las sondas para el ARN *APETALA3* (*amarillo*) se unen a las regiones que se diferenciarán como pétalos y estambres.

La clonación de estos genes de identidad de órganos no sólo posibilitó los experimentos de hibridación *in situ*, sino que produjo también otros beneficios de mayor alcance. Al analizar el ADN de los genes, empezamos a conocer las proteínas que determinaban y observar su intervención en el nivel subcelular.

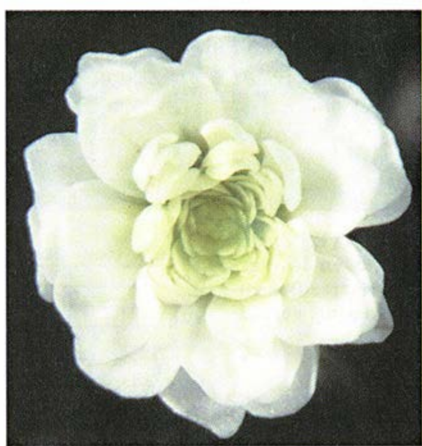
Así, buena parte de la proteína codificada por el gen *AGAMOUS* es muy parecida, en su secuencia de aminoácidos, a ciertas proteínas de levaduras, seres humanos y otros organismos. Se sabe que este tipo de proteínas se une al ADN y regula la transcripción de los genes en ARN. En otras palabras, estas proteínas actúan a modo de conmutadores moleculares, activando o desactivando genes mediante su unión a secuencias del ADN próximas.

El equipo de A. Nordheim, de la Facultad de Medicina de Hannover, ha demostrado que la proteína de *AGAMOUS* se une efectivamente al ADN. Es, por tanto, muy verosímil que, como otras proteínas del mismo tipo, desempeñe una función reguladora.

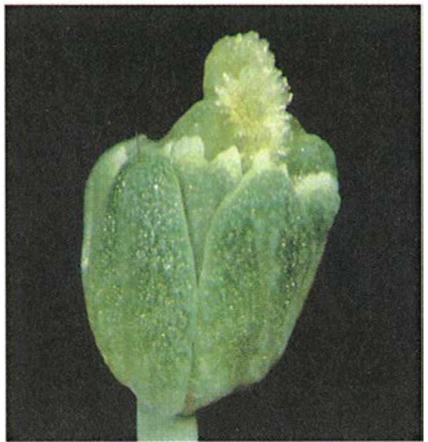
No sabemos a qué genes se une, y regula, la proteína *AGAMOUS* en la flor de *Arabidopsis*. Pero podemos suponer que controla la actividad de una serie de genes que preceden en la secuencia. Genes que, a su vez, pueden ser reguladores de otros genes o fabrican quizás el producto proteínico final que distingue un tipo de célula floral de otro. (A modo de inciso: hay un pariente humano de *AGAMOUS* que determina el factor de respuesta serológica, proteína que regula el oncogén *c-fos*, implicado posiblemente en cánceres humanos. Este notable paralelismo entre control genético de plantas y de seres humanos subraya la unidad evolutiva de la vida.)

Mientras clonábamos *AGAMOUS* en *Arabidopsis*, Hans Sommer, Zsuzsanna Schwarz-Sommer y Heinz Saedler, del Instituto Max Planck de Mejora Vegetal en Colonia, hacían lo propio con el gen *DEFICIENS* en flores de boca de dragón (*Antirrhinum*). *DEFICIENS* es el equivalente a la actividad B en estas plantas (muy remotamente emparentadas con *Arabidopsis*). Los mutantes en *DEFICIENS* se conocen desde hace décadas. Significativamente, *DEFICIENS* produce la misma clase de proteínas que se unen al ADN a la que pertenece *AGAMOUS*, lo que sugería que podría ocurrir lo mismo con otros genes relacionados con la forma floral.









9. LOS GENES *SUPERMAN* y *LEAFY* intervienen en una fase precoz del desarrollo de la flor y deciden qué células expresarán las actividades A, B y C. *SUPERMAN* regula los genes de la actividad B. De este modo, los mutantes *superman* (*arriba*) muestran estambres en el verticilo 4 y una correspondiente reducción en los carpelos. Los dobles mutantes *superman agamous* (*segundos por arriba*) tienen pétalos en los verticilos 2, 3 y 4. Los mutantes *superman agamous apetala2* (*terceros por arriba*) presentan hojas en torno a unos órganos intermedios entre estambres y pétalos. Las flores mutantes *leafy* (*abajo*) sólo tienen órganos parecidos a sépalos dispuestos en espiral, como las hojas en los tallos.

El grupo alemán clonó luego una serie de genes reguladores de *Antirrhinum*. Hasta ahora, los genes de *Antirrhinum* corresponden punto por punto a los de *Arabidopsis*, tanto en las secuencias de las proteínas correspondientes que codifican como en los tipos de anormalidades en la forma que resultan de su pérdida. Nuestros colegas alemanes y nosotros hemos encontrado en ambas plantas muchos miembros de esta familia de genes que no están correlacionados con ningún mutante floral, si bien se hallan activos en regiones de las flores en desarrollo. Esta familia de genes podría implicar otras claves adicionales en el desarrollo floral.

Hemos arribado así a una situación en que un grupo escueto de genes reguladores determinan la identidad de un órgano volviéndose activos de una manera altamente específica y controlando, a su vez, otros genes en una cascada reguladora. Pero, ¿en virtud de qué esos genes saben cómo activarse en el lugar correcto y en fases muy tempranas del desarrollo floral?

La respuesta es que estos genes se alojan en la mitad, no al principio, de la cascada reguladora. Las células que expresan genes de identidad de un órgano deben ser distintas de sus vecinas, incluso antes de que estos genes se activen. Las regiones concéntricas que más adelante representarán el dominio del destino de los órganos deben estar ya establecidas. Por tanto, los genes de identidad de los órganos han de suponer un paso esencial en la *interpretación* intracelular de la posición, aunque ellos mismos no estén involucrados en el *establecimiento* inicial de las diferencias posicionales.

Otro grupo de genes reguladores, que actúan en una fase precoz del desarrollo, parece acotar los dominios posicionales dentro de las flores. Uno de estos genes recibe el nombre

de *SUPERMAN* (estudiado con el nombre *FLO10* por George W. Haughn). En las flores de las plantas a las que falta este gen, el cuarto verticilo no contiene carpelos, sino más estambres (por esa razón se denominó *superman* al mutante). De acuerdo con el modelo genético, esta disposición implica que la actividad B interviene en el cuarto verticilo de las flores mutantes, lo cual sugiere a su vez que el gen *SUPERMAN*, en su operación normal, impide la actividad B en ese verticilo.

Si es así, podemos entonces proponer la hipótesis de que las flores con mutaciones en los dos genes, *SUPERMAN* y *APETALA3*, se comportarán como las mutantes *apetala3*; vale decir, si *APETALA3* no es funcional, carece de significado que esté incorrectamente regulado. Los experimentos de hibridación han confirmado tal hipótesis; y, parejamente, corroboraron otras dos: que los mutantes a los que faltan las dos actividades *SUPERMAN* y *AGAMOUS* tendrán sépalos en torno a tres verticilos de pétalos y que los mutantes triples de la variedad *superman agamous apetala2* tendrán hojas alrededor de tres verticilos de órganos intermedios entre estambres y pétalos.

Los experimentos de hibridación *in situ* respaldan también la idea de que *SUPERMAN* regula los genes de la actividad B. El ARN de *APETALA3* y *PISTILLATA* aparece en el cuarto verticilo de los mutantes *superman*, así como en las posiciones habituales de los verticilos segundo y tercero.

La función del gen *SUPERMAN* explica por qué los genes de la actividad B no se activan en el cuarto verticilo. De todos modos, no explica por qué esos genes se activan en el segundo y tercer verticilo en el momento adecuado del desarrollo. La respuesta parece implicar un gen que actúa incluso antes, *LEAFY* [FOLIOSO]. Las flores de las plantas que carecen de un gen *LEAFY* activo carecen de pétalos y estambres. Presentan, en su lugar, órganos sepaloides. Pero estos órganos, en vez de crecer en anillos concéntricos, a la manera de los órganos normales de los verticilos segundo y tercero, se disponen en espiral, a semejanza de las hojas en el tallo de una planta de *Arabidopsis*.

El gen *LEAFY* comunica, según parece, a las células del segundo y tercer verticilos que se encuentran en una flor; sin ese aviso, las células se dividen con una disposición que es habitual en el vástago, fuera de las flores. Por su aspecto sepaloide, podrá inferirse que en tales órganos se desempeña la actividad A y que no es necesaria la presencia de *LEAFY* para la activación inicial de los genes A. Tampoco parece requerirse *LEAFY* para la iniciación de la actividad C, pues puede haber carpelos en el centro de las flores mutantes *leafy*.

Sin embargo, *LEAFY* parece ser imprescindible para la expresión inicial de la actividad B, porque sin él no se forman pétalos ni estambres. También debe ser la clave de funciones más fundamentales que especifiquen la posición relativa y el orden radial de los órganos.

Los resultados de las pruebas genéticas y moleculares son coherentes con esa conclusión. Los mutantes que carecen de función en los dos genes *LEAFY* y *APETALA3* muestran el mismo aspecto de aquellos a los que sólo les falta *LEAFY*: ninguno tiene actividad B. También, en los mutantes *leafy*, la cantidad de ARN de *APETALA3* es mucho menor que en las flores normales. Por consiguiente, la proteína producida por *LEAFY* es probablemente un activador de *APETALA3*.

Los experimentos realizados para descubrir cuándo se activa el gen *LEAFY* han dado resultados coherentes con nuestros modelos genéticos originales. El grupo de Enrico S. Coen, del Instituto John Innes de Norwich, ha clonado un gen con una función parecida a la de *LEAFY* en un *Antirrhinum*. Usando ese gen como sonda, clonamos luego nosotros *LEAFY* de *Arabidopsis*. A tenor de la investigación sobre ARN, *LEAFY* y el gen emparentado de *Antirrhinum* parecen ser los primeros genes en intervenir en el desarrollo floral. Sus productos de ARN emergen cuando se forman los grupos iniciales de células que se convertirán más tarde en primordios florales, mucho antes de que se tornen activos los genes de identidad de órganos.

Las mutaciones en otros genes indican que éstos cumplen también funciones muy tempranas en el desarrollo floral. Además, no nos cabe la menor duda de que se descubrirán nuevos genes de identidad de órganos. Sea como sea, contamos ya, como mínimo, con un esquema de las jerarquías genéticas que controlan el desarrollo de la flor en *Arabidopsis*. Según han revelado los grupos del Instituto John Innes y del Max Planck, los genes de *Arabidopsis* pueden asimilarse a los que se conocen en *Antirrhinum*. Genes emparentados dirigen la creación de forma en el desarrollo de otras flores, confirman investigaciones recientes. En consecuencia, el modelo genético del desarrollo floral que hemos pergeñado pudiera resultar válido para muchas flores de disposición radial, si no para todas ellas.

Equipados con un modelo genético predictivo y conocedores de algunos de los genes que especifican los órganos florales, los investigadores han alcanzado un notable grado de control sobre el desarrollo de las flores. Se pueden extraer los genes clonados, alterar las proteínas que codifican o sus modos de actuación y reintroducirlos en las plantas para modificar a voluntad el desarrollo floral.

Hay un proyecto de colaboración entre Martin F. Yanofsky, antiguo alumno mío hoy en la Universidad de California en San Diego, y nuestro laboratorio que podría abrir el camino para los experimentos rutinarios de mañana. Trabajando con clones del gen *AGAMOUS*, alteramos las secuencias de control anejas que regulan el momento y el lugar de su activación. Sustituimos estas secuencias con otras procedentes de un gen diferente que se expresa en todas las células. Luego, insertamos el gen recombinante en plantas de tabaco. De acuerdo con lo esperado, las plantas resultantes tenían carpelos (y óvulos en desarrollo) en las posiciones que normalmente deberían ocupar los sépalos, y estambres allí donde correspondían pétalos.

No se saca ninguna utilidad directa de estas plantas de tabaco manipuladas. Pero constituyen una prueba viva del poder que, sobre la estructura y función de algunas plantas, nos ha concedido el conocer mejor las bases moleculares del desarrollo floral.

BIBLIOGRAFIA COMPLEMENTARIA

ARABIDOPSIS THALIANA AND PLANT MOLECULAR GENETICS. E. M. Meyerowitz y R. E. Pruitt en *Science*, vol. 229, págs. 1214-1218; 20 de septiembre de 1985.

ABNORMAL FLOWERS AND PATTERN FORMATION IN FLORAL DEVELOPMENT. E. M. Meyerowitz, D. R. Smyth and J. L. Bowman en *Development*, vol. 106, n.º 2, páginas 209-217; junio de 1989.

GENETIC CONTROL OF FLOWER DEVELOPMENT BY HOMEOTIC GENES IN *ANTIRRHINUM MAJUS*. Z. Schwarz-Sommer, P. Huijser, W. Nacken, H. Sædler y H. Sommer en *Science*, vol. 250, páginas 931-936; 16 de noviembre de 1990.

THE WAR OF THE WHORLS: GENETIC INTERACTIONS CONTROLLING FLOWER DEVELOPMENT. Enrico S. Coen y E. M. Meyerowitz en *Nature*, vol. 353, n.º 6339, páginas 31-37; 5 de septiembre de 1991.

Marguerite Holloway

Mary Leakey: *Desterrar la historia*

Mary Leakey contesta, rápida y breve, mis preguntas protegida por la cortina de humo que teje su puro holandés. Ni la arqueología ni la antropología estarían hoy donde están sin las excavaciones, pulcras y tesoneras, de esta anciana venerable. En una profesión minada por la lucubración interpretativa y la ambición personalista, Leakey destaca por su reticencia a especular sobre las teorías de la evolución del hombre.

Mary llegó a Kenia y Tanzania en 1935 con su esposo, el paleontólogo Louis Leakey, y salvo breves escapadas a Europa y Estados Unidos, ha permanecido allí desde entonces. Durante todos estos largos años, introdujo las técnicas de la arqueología moderna en el trabajo de campo realizado en Africa, y, usándolas, desenterró útiles de piedra y restos de humanos primitivos que han revolucionado la interpretación de nuestros orígenes. Sus descubrimientos han hecho que *Proconsul*, la garganta de Olduvai, el cráneo de *Zinjanthropus* y las huellas de Laetoli constituyan referencias habituales en los libros de texto.

Nació en Inglaterra y se crió en Francia. Independiente, puntillosa y anticonvencional desde siempre, visitó con su padre las pinturas prehistóricas de Fond de Gaume y La Mouthe; él le introdujo en las herramientas de piedra y de hueso que estudiaban los prehistoriadores franceses. Estos trabajos de arte le predispusieron a la excavación, el dibujo y la prehistoria: "Para mí representaba la diáfana alegría instintiva de coleccionar, o hasta podría decirse que de buscar tesoros; parecía como si esa región rezumara objetos de belleza e interés intrínseco que podían sacarse del suelo."

Esa educación acabó induciendo a Leakey a trabajar, a la edad de 17 años, en expediciones arqueológicas por Gran Bretaña. También asistió a clases de arqueología, prehistoria y geología en el Museo de Londres y en el Colegio Universitario de la capital. Reconoce que le faltaba paciencia para seguir los cursos académicos. Lo que no impide que se sienta satisfecha de sus múltiples distinciones honoríficas: "Al fin y al cabo, me las gané cavando bajo un sol de justicia."



La arqueóloga Mary Leakey posa con Sam, su dálmata

Cierta tarde, al acabar la clase, se topó con Louis Leakey. En 1934, el afamado investigador le pidió a Mary, ya conocida por su talento artístico, que ilustrara un libro suyo. Pronto estaban camino de Africa oriental. Formaban un equipo extraordinario. "Lo más destacable de mi madre es que no es muy de figurar y que trabaja de lo lindo", dice Richard E. Leakey, un iconoclasta abogado de la prohibición del tráfico de marfil y reputado paleontólogo. "Su preocupación por los detalles y la búsqueda de la perfección labraron la carrera de mi padre. El solo nunca hubiera sido famoso. Ella era mucho más ordenada, algo así como un técnico. El, con talante de mago, se excitaba con mayor facilidad."

Lo que la mente exigente y el mago encontraron en los años que dedicaron a desenterrar el pasado no se consiguió sin esfuerzo y penuria. Desde 1935 hasta 1959 los dos trabajaron en diversos puntos de Kenia y Tanzania, buscando los esquivos restos de los primeros humanos. Hubieron de hacer frente a obstáculos de toda laya, desde las duras condiciones del entorno hasta la escasa financiación. Los éxitos también fueron magros hasta 1948, año en que Mary encontró el primer cráneo perfectamente conservado y los huesos faciales de un hominoide, *Proconsul*, de unos 16 millones de años. Este antropomorfo del Mioceno, hallado en la isla Rusinga, del lago Victoria, suministró a la arqueología el primer cráneo de lo que se creía era el eslabón perdido, un mono arborícola dotado de un cerebro mayor que el de sus predecesores.

Proconsul fue un hallazgo maravilloso, pero no trajo más dinero. Los Leakey siguieron en sus duras condiciones hasta 1959. Cierta mañana de ese año, en la garganta de Olduvai, zona de Tanzania cercana al Gran Valle del Rift que cruza Africa oriental de norte a sur, salió Mary a ver qué encontraba. De una de las secciones excavadas asomaba un cráneo de homínido de unos 1,8 millones de años, al poco apodado *Zinjanthropus*. Zinj fue el primero de un nuevo grupo (*Australopithecus boisei*) y el primero de

esos cráneos que se encontró en Africa oriental.

"Por alguna razón, aquel cráneo disparó la imaginación", recuerda. "Pero también despertó, y eso fue muy importante para nosotros, el interés de la National Geographic Society, que nos subvencionó durante bastantes años."

Leakey no especulará sobre la ubicación de Zinj en el árbol genealógico. "Nunca tuve la sensación de que la interpretación fuera lo mío. Yo vine a cavar y a sacar las cosas lo mejor que pudiera", asevera con plana sencillez. "No sabemos mucho, y cuanto más sabemos más nos damos cuenta de cuán erradas iban las primeras interpretaciones. La interpretación es un ejercicio mental necesario, pero la gente se acalora mucho y pierde la compostura, cayendo en el ridículo."

Trato de sonsacarle algo sobre otra cuestión candente: ¿Apareció *Homo sapiens*, en Africa? ¿Procede acaso de ancestros dispersos por doquier, como dice la hipótesis multirregional? Leakey esboza una sonrisa. "No va a obtener de mí nada sensacional sobre esas disputas. Si yo fuera Richard, le hablaría durante horas y horas, pero no me parece que valga la pena." Se calla. "Lo mío es pisar terreno firme, y ése nunca lo es."

En el campo, Leakey se hallaba a sus anchas, sobre tierra firme. Los yacimientos fueron cuidadosamente cartografiados y datados, y su estratigrafía, rigurosamente anotada. Además de los restos de homínidos encontrados y catalogados en Olduvai, Mary descubrió herramientas de hasta dos millones de años de antigüedad: útiles de piedra del período Oldovano. También registró la evolución de los artefactos en el curso del tiempo, estableciendo una segunda forma, Oldovano avanzado, que perduró hasta hace unos 500.000 años.

"La comunidad arqueológica debería estar agradecida de que ella estuviera a cargo de Olduvai", afirma Rick Potts, antropólogo de la Institución Smithsoniana que está estudiando Olorgesailie, un yacimiento al sur de Nairobi donde los Leakey encontraron hachas de piedra en 1942. Ahora, como entonces, la blanca y arenosa sabana masai estaba abarrotada de útiles. Los más bellos han sido robados. Leakey se alegra de que la Institución Smithsoniana restaure el yacimiento y su pequeño museo, y esté decidida a acotar la zona.

Suerte que no disfruta la garganta de Olduvai. Después de años de vivir y trabajar allí, y tras la muerte de Louis en 1972, Mary se retiró en 1984. Desde entonces, ha trabajado en su obra sobre los descubrimientos de Olduvai y ha escrito un libro sobre las pinturas rupestres de Tanzania. "Me hice muy vieja para vivir al raso", explica. "Se necesitan buena salud y arrestos juveniles. No tenía sentido que me empecinara." Tras su marcha, el yacimiento quedó abandonado. "Voy una vez al año a Serengeti a ver la migración de los animales salvajes, porque eso significa mucho para mí, pero evito Olduvai siempre que puedo. Aquello es una ruina. Me deprime." Con voz airada, entona una letanía de pesares: la excavación abandonada, el museo cayéndose, los artefactos robados, los catálogos perdidos. "Afortunadamente, todavía queda mucho bajo tierra. Es un lugar riquísimo, y hay mucho más bajo la superficie para las futuras generaciones que estén mejor formadas."

Del descubrimiento más espectacular de Leakey, y el que ella considera más importante, realizado en 1978, no queda nada; fue destruido cuando dejó el lugar. Las huellas de Laetoli, una zona cercana a Olduvai, proporcionaron al mundo la primera prueba de la marcha bípeda. Tres homínidos habían caminado sobre la arena volcánica, que se fosilizó, conservando las pisadas. El suelo tenía unos 3,6 millones de años. Aunque algunos huesos fosilizados de extremidades inferiores de otros homínidos ya lo sugerían, el descubrimiento marcó la edad del bipedismo.

"La interpretación es un ejercicio mental necesario, pero la gente se acalora mucho y pierde la compostura, cayendo en el ridículo."

Hoy las famosas huellas sólo podrían ser rescatadas con la intervención del Instituto Getty para la Conservación. "Oh, se encuentran en un estado terrible", exclama Leakey. "Cuando me marché, las cubrí con una capa de arena de río, plásticos, más arena y piedras encima para protegerlas de los animales y de los masai". Pero han crecido acacias entre las huellas y sus raíces las han destrozado.

Aunque se ha mantenido al margen de la controversia en sus escritos, no ha escapado del todo. Ella y Donald Johanson, paleontólogo del Instituto del Origen del Hombre de Berkeley, polemizaron sobre la relación entre los homínidos encontrados en Etiopía y en Laetoli. Se puso en discusión el número de pisadas que había en Laetoli. Tim White, de la Universidad de California en Berkeley, sostenía que sólo había dos y que Leakey y su equipo habían hecho la otra con una herramienta durante la excavación. ¿La respuesta de Leakey? "Vaya una tontería", se ríe y da carpetazo al asunto.

Hay un tema que aborrece. "¿Cómo le fue, siendo mujer? ¿Y esposa? ¿Y madre?" Responde: "Necedades." A Leakey (al igual que muchas otras mujeres de ciencia de su generación, como Rita Levi-Montalcini y Gertrude Belle Elion, ambas galardonadas con el Nobel) no le agradan las preguntas que plantean qué hace una mujer en el terreno de los hombres. Su sexo no tuvo relevancia alguna en el trabajo, asegura. Sólo hizo lo que quería hacer. "Nunca fui consciente de ello. No estoy diciendo una mentira en aras de nada. Nunca me encontré en inferioridad de condiciones."

Leakey se limitó a trabajar, por encima de discusiones profesionales y revueltas políticas. En 1952, durante el alzamiento Mau Mau, pusieron precio a la vida de Louis, que había sido recibido en la tribu Kikuyu durante su niñez en Africa. Los cuatro años que duró la rebelión fueron terribles para el país. Sabotearon los frenos del coche de Mary y un pariente de Louis fue asesinado. La casa donde Leakey vive ahora fue proyectada en aquella época: una estructura blanca, baja, con un patio central por donde los perros corretean por la noche. Esos cachorros son muy importantes para Leakey: suponen una compañía y brindan seguridad en la maleza.

Parece que hemos tocado todos los temas, de modo que me hace un resumen de sus descubrimientos en voz alta. "Pero no ha mencionado usted los frutos", recuerda. Uno de los hallazgos favoritos de Leakey es un surtido de fósiles del Mioceno: frutos enteros, semillas, insectos (todo un nido de hormigas, entre ellos) y un lagarto con la lengua fuera. Se hallaban en el terreno arenoso de la isla Rusinga. "Los encontramos simplemente porque nos sentamos a fumar un cigarrillo, calurosos y cansados, los vimos sobre el suelo a nuestro alrededor. Antes habíamos pasado caminando sobre ellos por todo el lugar." Se detiene. "Sabe, uno en realidad sólo encuentra lo que está buscando."

Ciencia y sociedad

Desarrollo embrionario

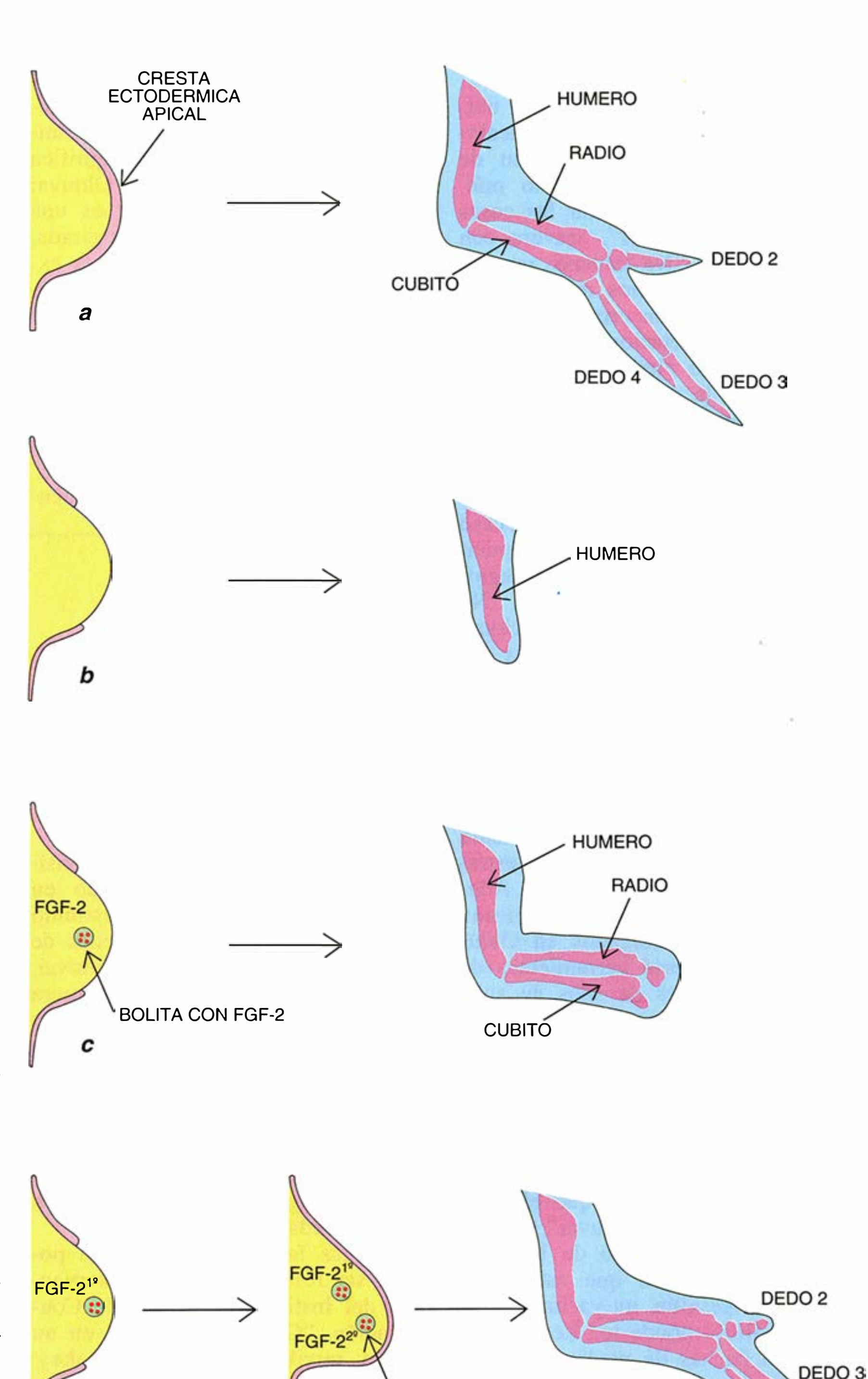
Factores de crecimiento fibroblástico

El desarrollo de los organismos plantea numerosas preguntas al biólogo. Por ejemplo: ¿cómo logran las células organizarse y formar una determinada estructura? Ha de haber señales que indiquen a las células su posición en el sistema y éstas deben saber interpretarlas y comportarse de acuerdo con las exigencias de dicha asignación.

Entre los sistemas de desarrollo mejor conocidos se encuentra la formación de las extremidades de los vertebrados. Ello obedece a una doble razón; en primer lugar, su fácil acceso para la manipulación experimental, especialmente *in ovo* en el embrión de pollo; en segundo lugar, porque se trata de un sistema en el que, desde el comienzo, existe toda la información necesaria para su desarrollo, aun cuando se aísle del resto del embrión.

El desarrollo del organismo arranca de tres hojas embrionarias: endodermo, mesodermo y ectodermo. El esbozo embrionario que da lugar a una extremidad consta de una masa de células mesodérmicas, recubiertas por ectodermo (que originará la epidermis de la piel). Durante el curso del desarrollo, las células mesodérmicas se organizan y se diferencian para constituir los diversos elementos componentes de la extremidad adulta (esqueleto, musculatura, tendones, etcétera).

Sabemos que las señales necesarias para que se desencadenen estos procesos parten de grupos celulares especializados que el esbozo posee desde los estadios iniciales de su desarrollo. Así, el ectodermo más distal de la punta de la extremidad, llamado la "cresta ectodérmica apical", es un elemento crucial responsable del crecimiento próximo-distal (del tronco hacia los dedos) de la extremidad. La extirpación quirúrgica de la cresta comporta el cese del crecimiento y consiguiente formación de extremidades truncadas en las que no se desarrolla la parte más distal. Cuanto más temprana sea la fase del desarrollo en que se elimine la cresta, más próximo al tronco se hallará el



Efecto de la extirpación de la cresta ectodérmica apical durante el desarrollo de la extremidad de pollo, con y sin administración de FGF-2 exógeno: (a) *desarrollo normal;* (b) *desarrollo después de extirpar la cresta;* (c) *desarrollo tras extirpar la cresta y administrar una bolita de gel impregnada con FGF-2, y* (d) *desarrollo tras la extirpación de la cresta y administración secuencial de dos bolitas de gel impregnadas con FGF-2. Nótese en el dibujo que la herida dejada en el ectodermo al quitar la cresta cura, pero no se regenera una nueva cresta ectodérmica apical*

truncamiento de la extremidad resultante. Vemos, además, que los esbozos de extremidad dotados de dos crestas ectodérmicas apicales, como observamos en una eudiplopodia por mutación o en el caso de un injerto microquirúrgico de una cresta extra, dan lugar a la formación de extremidades duplicadas. De acuerdo con la hipótesis vigente, la cresta produce un factor que mantiene las células mesodérmicas subyacentes en un estado de rápida proliferación e indiferenciación, constituyendo la "zona de progreso". Una vez que las células se alejan de ésta, en virtud de las múltiples divisiones, se diferencian en los elementos esqueléticos siguiendo un orden próximo-distal.

El tiempo que las células permanecen bajo la influencia inmediata de la cresta parece ser un factor clave en la asignación de su misión o destino en el eje próximo-distal: las células que han estado menos tiempo bajo la influencia de la cresta ectodérmica apical desarrollan el húmero/fémur, mientras que las que han estado más tiempo se ven determinadas a formar los dedos.

Nuestro trabajo se ha centrado en la identificación de las bases moleculares implicadas en la relación entre la cresta ectodérmica apical y el mesodermo subyacente. Para ello, utilizamos el modelo de desarrollo de la extremidad del embrión de pollo. Los estudios en cultivo realizados con anterioridad apuntaban hacia la familia de los factores de crecimiento fibroblástico (FGF), integrada por nueve miembros. En particular, el FGF-2 parecía el candidato mejor colocado; habíase confirmado su presencia en el embrión de pollo a lo largo de experimentos de inmunohistoquímica donde se empleaban anticuerpos específicos contra el FGF-2. Además, su pauta de distribución en la extremidad era compatible con una posible función señalizadora de la cresta.

Nos aprestamos, pues, a sustituir el efecto de la cresta ectodérmica apical por FGF-2 administrado exógenamente, en un bolita de gel. Al tercer día del desarrollo, extirpamos la cresta de la extremidad, e implantamos inmediatamente en la zona de progreso una bolita de gel impregnada con FGF-2. ¿Qué aconteció? Los embriones control (cuya bolita de gel estaba impregnada sólo con tampón) presentaban extremidades truncadas en el punto esperado (húmero); por contra, los embriones con FGF-2 continuaban creciendo y el nivel de truncamiento de la extemidad era mucho más distal (en la muñeca).

Examinamos la cinética de liberación del FGF-2 a partir de la bolita de gel. Se comprobó que la mayor parte del FGF-2 se liberaba durante las primeras 24 horas. Procedimos entonces a la implantación de una segunda bolita de FGF-2 en la zona de progreso, 24 horas después de la implantación de la primera, lo que nos permitía conseguir niveles continuados de FGF-2 en la extremidad. El 100 % de los embriones que, tras serles extirpada la cresta, recibieron secuencialmente dos bolitas con FGF-2 desarrollaron una extremidad completa. Así se confirmó la capacidad del FGF-2 para sustituir a la cresta ectodérmica apical.

El FGF-2 se mostró también capaz de mantener otros aspectos celulares y moleculares que dependen de la presencia de la cresta apical; por ejemplo, la viabilidad de las células de la zona de progreso y una expresión génica característica de esta región. La extirpación de la cresta provoca la muerte en masa de las células de la zona de avance, mortandad que se evita con la implantación del FGF-2 exógeno. Sabemos, asimismo, que la expresión, por las células de la zona de progreso, del gen homeótico Msx1 está regulada por la cresta. La expresión de este gen decae bruscamente hasta tornarse indetectable a las tres horas de la extirpación de la cresta; pero la administración de FGF-2 logra mantener dicha expresión en sus niveles de normalidad.

El grupo de Lee Niswander ha demostrado que otro miembro de la familia de los FGF, el FGF-4, es capaz, lo mismo que el FGF-2, de sustituir la acción de la cresta cuando se le administra exógenamente tras la extirpación de ésta durante el desarrollo de la extremidad.

Para dilucidar cuál de estos dos factores era la señal que emana de la cresta *in vivo* se recurrió a un bioensayo basado en la capacidad de los FGF de reprimir la diferenciación de una línea de mioblastos de ratón (MM14). El bioensayo en cuestión permite detectar cantidades mínimas de FGF ($0{,}3 \times 10^{-15}$ mol/mililitro). Se prepararon extractos de las células del esbozo de extremidad y se comprobó su capacidad para reprimir la diferenciación de las células MM14, lo que indicaba la presencia de actividad FGF en los extractos. Para poder distinguir si este efecto se debía al FGF-2 o al FGF-4, se empleó un anticuerpo específico contra FGF-2, que bloqueaba la actividad del FGF-2 pero no la del FGF-4. Y se observó que, al añadir el anticuerpo a los extractos del esbozo,

desaparecía prácticamente toda su actividad sobre los mioblastos MM14.

Suelen aceptarse tres criterios para establecer que una sustancia constituye la señal utilizada para la comunicación entre células o tejidos. De acuerdo con el primer criterio, su localización ha de ser la adecuada; a tenor del segundo, dicha sustancia ha de poder remplazar, en todos los niveles, la presencia de las células que la producen; por último, dicta el tercer criterio que el bloqueo de esta sustancia provoque la pérdida de la señal. Nosotros hemos demostrado el cumplimiento de los dos primeros criterios por el FGF-2. Nos falta corroborar el tercero.

Los datos aquí expuestos hacen altamente verosímil que el FGF-2 sea el factor de crecimiento producido por la cresta ectodérmica apical en el embrión de pollo y, por tanto, la base molecular que controla el crecimiento de la extremidad durante su desarrollo.

María A. Ros
Dpto. de Anatomía y Biología Celular. Universidad de Cantabria

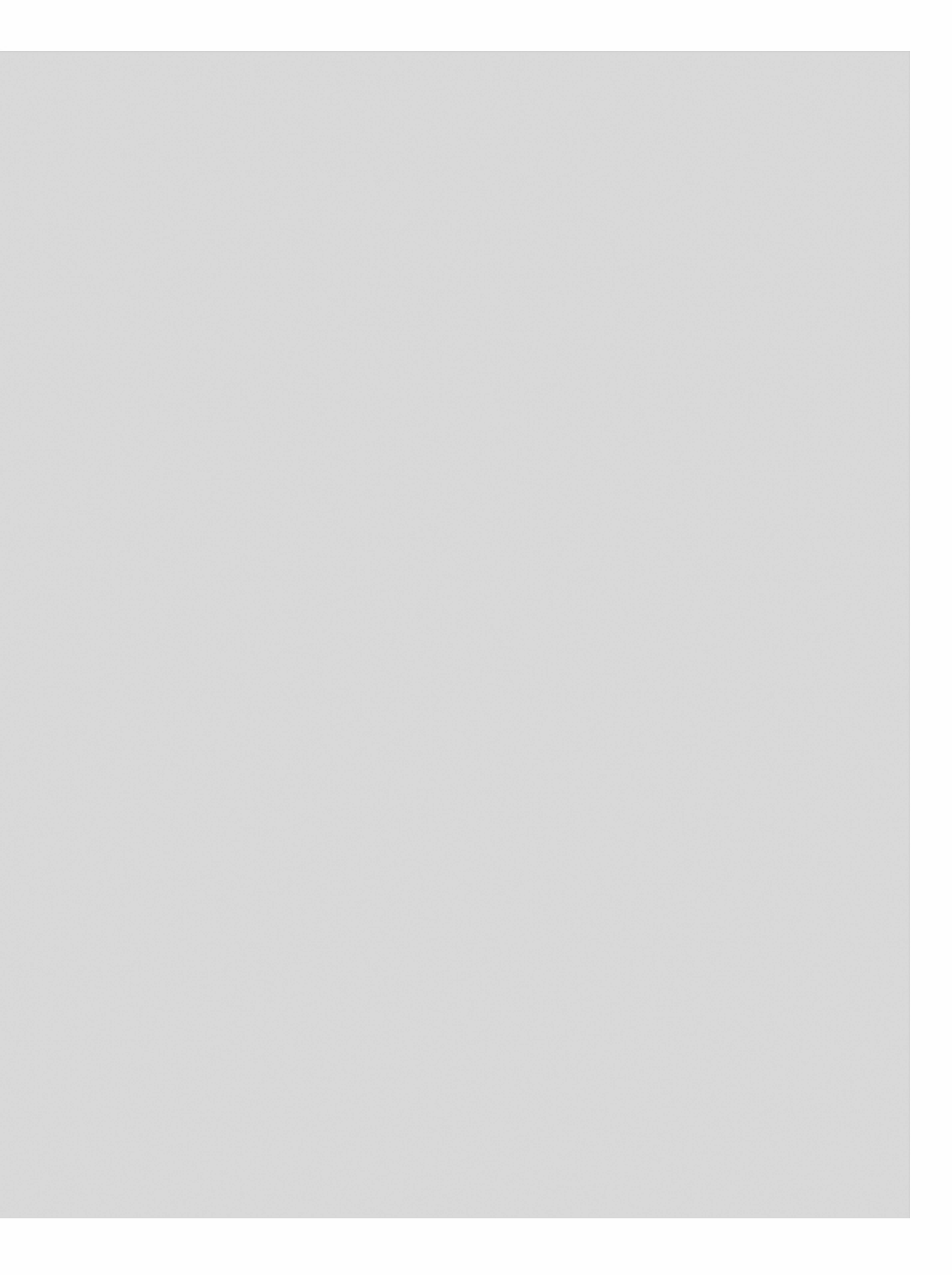
John F. Fallon
Dpto. de Anatomía.
Universidad de Wisconsin

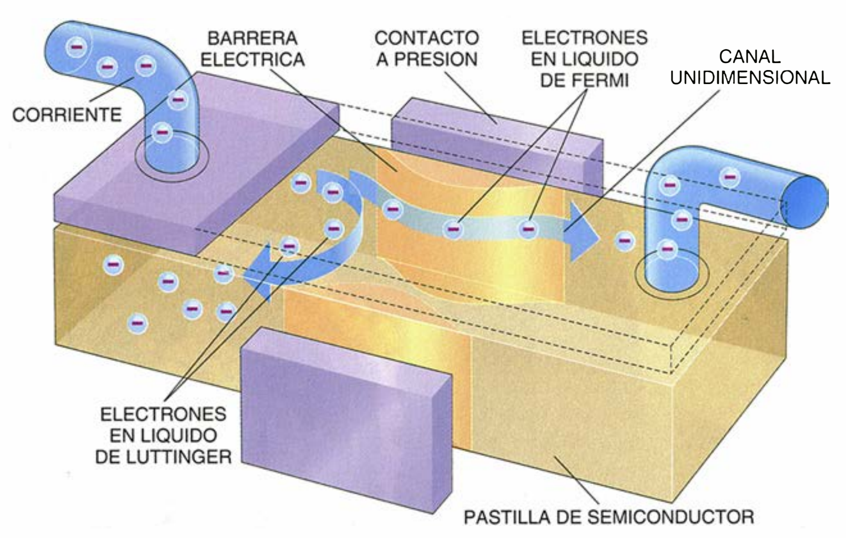
El problema de "muchos cuerpos"

Simplemente una fase

Cuando se nos pide describir las diversas fases de la materia, lo primero que nos viene a la cabeza es el agua, el hielo y el vapor. Para un físico, no obstante, cualquier configuración de partículas o entidades distinta de lo habitual puede también constituir un nuevo estado; ocurre así, por ejemplo, cuando los electrones se organizan en una onda de densidad de carga. Otra fase posible sería el líquido de Luttinger, que representa un comportamiento colectivo especial de los electrones en un medio conductor.

En condiciones normales, los electrones de un conductor forman lo que llamamos un líquido de Fermi. Este no es sino un mar de cargas negativas, dentro del cual un electrón cualquiera no responde a las cargas individuales de los demás electrones presentes en el material. El líquido de Fermi se compone de partículas que no interaccionan entre sí, lo que deja a los electrones vagar con entera libertad a través del cuerpo conductor. Esto en parte explica





Líquido de Luttinger formado en un canal grabado en una pastilla de semiconductor. En este estado los electrones se reflejan en una barrera eléctrica formada por un contacto a presión. Por el contrario, los electrones de un líquido de Fermi pueden salvar por efecto túnel el obstáculo

la manera en que los electrones transmiten la corriente por un conductor.

En los años sesenta, Joaquin M. Luttinger, de la Universidad de Columbia, investigó situaciones en las que los electrones se ven forzados a interaccionar entre sí. En un caso simplificado, unidimensional, resolvió las ecuaciones que definían ese estado: el problema de "muchos cuerpos". Así se quedaron las cosas hasta que los avances de la tecnología y el descubrimiento de la superconductividad en alta temperatura provocaron una nueva e intensa investigación de la actividad de los electrones en cuerpos sólidos.

En 1993 Charles L. Kane, de la Universidad de Pennsylvania, y Matthew P. A. Fisher, de la Universidad de California en Santa Bárbara, extrajeron de los cálculos de Luttinger una predicción susceptible de comprobación, y en la reunión de marzo de 1994 de la Sociedad Física de América, Webb, de la Universidad de Maryland, presentó la primera prueba experimental. Según Webb, la teoría señala el modo de constituir el sistema: los electrones han de circular por un canal unidimensional que pueda ser obstruido en un punto intermedio. Tal obstrucción puede conseguirse con un contacto que actúe de mordaza eléctrica. Basta con aplicar una tensión al contacto a presión para que el canal se estreche y se reduzca la conductancia.

En un líquido de Fermi, habría electrones que traspasarían por efecto túnel la obstrucción, y siempre quedaría alguna conductancia en el canal. No sería así, sin embargo, en un líquido de Luttinger. En tal estado, a temperaturas cercanas al cero absoluto, cada electrón acusaría las fuerzas procedentes de las cargas individuales de los demás electrones, y ello serviría para correlacionar sus comportamientos. El efecto de tal correlación sería una característica caída de la conductancia a través del contacto a presión; al final, todos los electrones acabarán reflejándose en la barrera.

Webb trabajó durante dos años en tal experimento, y en colaboración con Frank P. Miliken y Corey P. Umbach, del Centro de Investigación de IBM, consiguió crear un líquido de Luttinger en un semiconductor de arseniuro de galio. En teoría, el sello distintivo del líquido solamente se manifiesta en los electrones balísticos, es decir, electrones que se mueven en una misma dirección, sin dispersarse. La fuente de los electrones balísticos era el efecto Hall cuántico fraccionario, fenómeno que alude a la deriva lateral de los electrones conforme avanzan a través de una muestra sometida a un campo magnético externo. Como ya había observado Xiao-Gang Wen, del Instituto de Tecnología de Massachusetts, esos electrones se comportan de un modo balístico cuando se dan las condiciones adecuadas.

No es probable que el líquido de Luttinger encuentre aplicaciones, pues se desestabiliza a temperaturas que superen en más de un grado al cero absoluto. Tal vez su verdadero valor estribe en que, por fin, puede verse la interacción real de los electrones en un sólido. Los métodos clásicos de analizar los problemas "de muchos cuerpos" requieren combinar la intuición con un método de aproximación llamado teoría de la perturbación.

PHILIP YAM

Receptores

Mecanismo de activación

La proliferación celular es el resultado del balance entre la acción de estímulos inductores y la de inhibidores del crecimiento celular. Sin embargo, mientras que el estudio de las señales mitogénicas (inductoras de la división celular) ha merecido la atención de los investigadores durante un largo período de tiempo, era escaso el conocimiento de los elementos que mediaban las respuestas antiproliferativas.

Los estudios realizados durante los últimos diez años han consolidado la idea de que estos sistemas inhibidores del crecimiento son, en sus líneas generales, similares a los que promueven el crecimiento: ambos incluyen series de factores paracrinos con receptores específicos conectados a cascadas intracelulares de transducción de estas señales.

Muchos de estos factores paracrinos están constituidos por polipéptidos, entre los que destaca el factor de crecimiento transformante β, o TGF-β. Presenta éste una amplia multifuncionalidad, que depende del tipo celular donde ejerza su acción. Tal propiedad se ilustra en diversas manifestaciones: además de sus potentes efectos antiproliferativos en una amplia variedad de tipos celulares, TGF-β se halla involucrado en fenómenos muy dispares, que van desde la adhesión y migración celulares hasta el control de la producción de la matriz extracelular, pasando por procesos de desarrollo y reparación de tejidos.

La forma activa del TGF-β interacciona con una serie de receptores que se hallan en la superficie de las células diana. La investigación acerca de dicha interacción, en la que se recurrió a TGF-β marcado radiactivamente, puso de manifiesto que, en la mayoría de las líneas celulares estudiadas, había esencialmente tres tipos diferentes de moléculas con capacidad de unión del factor. En razón de su tamaño, se les denominó receptores de tipo I, II y III.

¿Por qué necesita la célula tres tipos diferentes de receptores para detectar la presencia de TGF-β en el medio extracelular? Se dio un paso

importante en el camino hacia la solución de esta cuestión con la obtención de células portadoras de mutaciones en alguno de los receptores, que no inhibían su crecimiento en presencia del factor. Estas células mutantes constituían, pues, un modelo idóneo para examinar la funcionalidad de los receptores a medida que se clonaban éstos.

Se pudo así observar que los receptores directamente involucrados en transmitir la señal inducida por el factor son los receptores de tipo I y II. Por sí solo, ninguno de estos dos receptores especifica respuesta alguna; se precisa la acción conjunta de ambos para todas y cada una de las respuestas celulares al factor. Por su parte, la caracterización del receptor de tipo III puso de manifiesto que es capaz de unirse con alta afinidad a TGF-β y "presentarlo" a los receptores I y II, reforzando de ese modo la acción ejercida por el factor cuando está presente a bajas concentraciones.

El análisis de la secuencia correspondiente a los receptores de tipo I y II reveló la semejanza de sus dominios intracelulares. Ambos receptores poseen, en esos dominios, estructuras con actividad quinasa. Es sabido que la fosforilación de las proteínas, catalizada por enzimas con actividad quinasa, constituye uno de los métodos habituales utilizados por los organismos para la transferencia de información. El examen ponía, pues, de manifiesto la posibilidad de que la información generada en la membrana plasmática se transformara en cambios en los niveles de fosforilación de proteínas. Hasta ese momento, se habían aislado numerosos receptores para factores de crecimiento y diferenciación con actividad quinasa dirigida específicamente a residuos de tirosina.

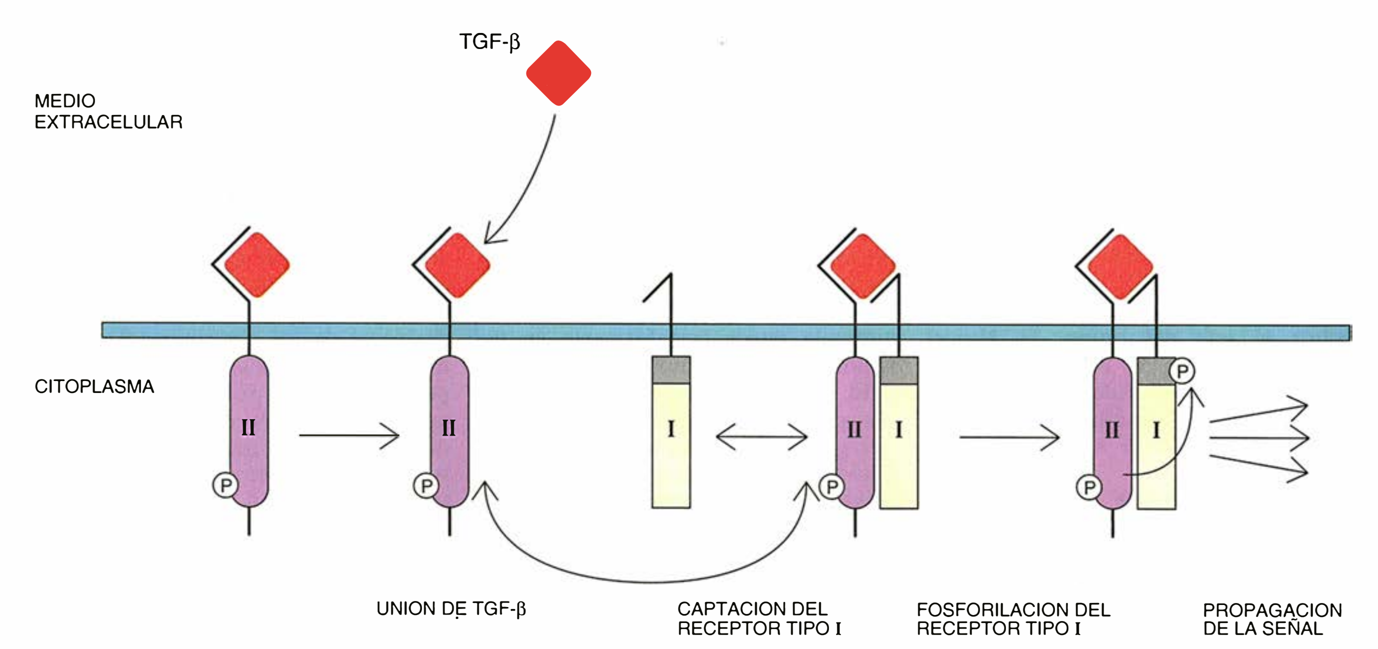
Para los receptores con actividad tirosina quinasa, el mecanismo de activación se genera a través de su dimerización, inducida por la presencia del factor. Este proceso promueve la autofosforilación de los receptores creándose lugares de interacción entre los receptores y los componentes que comunicarán la señal al resto de la célula.

En el caso de los receptores de TGF-β, la fosforilación está dirigida a otros aminoácidos: serina y treotina. Además, la necesidad de la acción conjunta de dos receptores diferentes, ambos con actividad quinasa (y siendo dicha actividad indispensable para la acción del TGF-β), sugería que el mecanismo de activación debía ser diferente de los receptores tirosinaquinasas. El receptor de tipo I es incapaz de unir por sí solo el factor y requiere la presencia del receptor de tipo II para ello. Una vez unido el factor, forman un complejo estable los tres componentes: TGF-β y receptores.

Investigaciones recientes han esclarecido este mecanismo. La unión de TGF-β al receptor de tipo II permite la captación del receptor de tipo I en un complejo ternario estable. La formación del complejo facilita la aproximación espacial de ambos dominios intracelulares, acercamiento necesario y suficiente para la fosforilación, por parte del receptor de tipo II, de una región concreta del receptor de tipo I, conocida como región GS, produciéndose así la activación de este último.

Queda mucho por averiguar sobre la propagación de las señales desde el complejo receptor activado. Ignoramos si esta fosforilación en la región GS puede producir una activación de la quinasa del receptor I o si, por contra, crea lugares de unión para otros componentes de la cascada de transmisión de la señal. Sin embargo, el hecho de que la actividad quinasa del receptor de tipo I, imprescindible para la acción del receptor, no participe en procesos de autofosforilación o no actúe sobre el receptor de tipo II, sugiere que probablemente sea esta actividad la implicada en la fosforilación y transferencia de la información a los otros componentes intracelulares de la vía.

En conclusión, de los resultados obtenidos hasta el momento emerge un modelo para la activación de estos receptores según el cual el factor se encarga directamente de poner en



Modelo para la activación del receptor de TGF-β. El primer paso en la activación corresponde a la unión de TGF-β (rombo) *por parte del receptor de tipo II. Esta unión induce la captación del receptor de tipo I en un complejo ternario estable, produciéndose activación del receptor de tipo I en virtud de la fosforilación de su región GS* (región punteada)

contacto dos receptores, ambos con actividad quinasa. La subsecuente fosforilación unidireccional por el receptor de tipo II del receptor de tipo I genera el primer eslabón en la propagación de las respuestas al factor. Las perspectivas futuras en este campo apuntan hacia la detallada caracterización de los mecanismos de activación y desensibilización del receptor y a la búsqueda de los sustratos del receptor activado.

FRANCESC VENTURA
Centro Oncológico Memorial
Sloan-Kettering, Nueva York

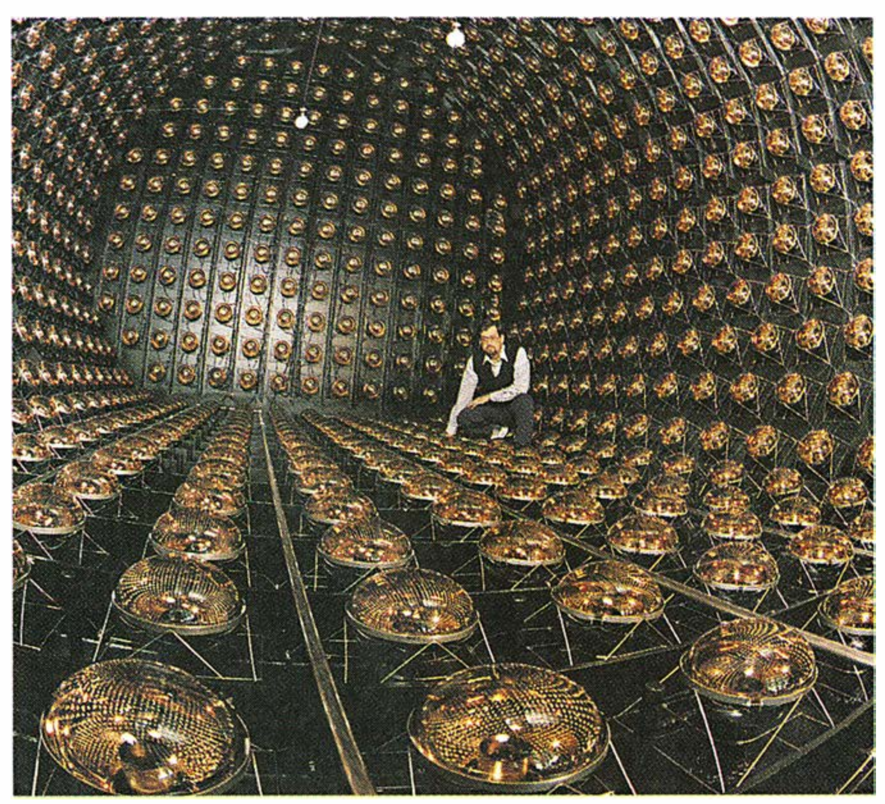
Caza de neutrinos

Con masa

Desde el día que descubrieron el neutrino, no han dejado los físicos de preguntarse si esta huidiza partícula, que se supone carece de masa, no tendrá en realidad alguna. Como en el universo hay un número muy grande de neutrinos, con que tuviesen una masa pequeña aportarían la "materia oscura" que, creen los cosmólogos, constituye la mayor parte de la sustancia del universo. Y, además, podrían convertirse en neutrinos de otros tipos mediante un proceso al que se le da el nombre de oscilación.

Un detector de partículas del Laboratorio Nacional de Los Alamos captó el año pasado ocho sucesos que podrían ser, quizá, las primeras oscilaciones de neutrinos que se hayan avistado. Si se confirman, estas observaciones probarían de paso que los neutrinos tienen masa. La sensibilidad del Detector de Neutrinos Líquido de Centelleo (DNLC) es la mayor del mundo por lo que se refiere a los cambios de los neutrinos. Pero quienes han participado en el experimento no hacen una proclamación explícita. "Nuestra impresión es que el peso de la prueba que recae sobre nosotros es grande", explica el jefe del grupo, D. Hywel White, "porque si es verdad, será muy importante". Además, los anuncios previos de oscilaciones neutrínicas han oscilado a su vez hasta desaparecer.

Hay tres tipos de neutrinos: electrónico, muónico y tau. ¿Por qué razón la capacidad que los neutrinos (o los antineutrinos) tienen de cambiar de un tipo a otro indica que tienen masa? La masa determina el ritmo con que la función de ondas de una partícula vibra. Si las ondas de dos neutrinos de masas diferentes se entremezclan, baten la una contra la otra de manera muy parecida a como lo hacen las ondas sonoras de distinto tono. Vemos a veces un neutrino, a veces el otro. Si los neutrinos carecieran de masa, sus ondas tendrían la misma frecuencia y no podrían batir en absoluto.



Captación de neutrinos mientras se cambiaban con los fototubos que forran este detector, que se muestra exento de líquido

No es fácil observar sucesos tan fugaces como la oscilación de los neutrinos. En el experimento del DNLC, se dispara un haz de protones procedente de un acelerador contra un blanco acuático. El blindaje que rodea este "amortiguador de haces" absorbe muchas de las partículas que se producen en las colisiones subsiguientes. Pero de vez en cuando un antineutrino muónico se escapa y vuela treinta metros hasta un detector lleno de aceite. Por el camino, la partícula puede transformarse en un antineutrino electrónico.

El antineutrino electrónico interactúa con un protón en el aceite, lo que genera un positrón y un neutrón. Pasado cierto tiempo, el neutrón se combina con otro protón y se emiten dos fotones, cuya energía total es característica. Los fotones y la traza brillante del positrón se observan mediante fototubos que forran el tanque del aceite. Para evitar la contaminación con otras partículas que puedan haberse infiltrado a través del blindaje, se buscan positrones que caigan dentro de un margen de energías específico.

El pequeño número de antineutrinos electrónicos observado sugiere que los antineutrinos muónicos sólo se convierten ligeramente en los electrónicos; el grado en que se mezclan es sólo un 1 %. Los experimentadores no citan la diferencia de masas. Pero los 30 metros a lo largo de los cuales los antineutrinos muónicos pueden cambiar de tipo quieren decir que el aparato es sensible a diferencias de masa de poco menos de un electronvolt. Esta diferencia implica una masa del neutrino grande.

Aunque el resultado podría servir para resolver el problema de la materia oscura, en nada contribuye a solventar otro de los rompecabezas con que disfrutan hoy los físicos: el problema de los neutrinos solares. El número de los neutrinos electrónicos procedentes del Sol es menos de la mitad del número que la teoría predice. El déficit podría explicarse si se aceptase que se convierten en neutrinos muónicos y, por tanto, eluden la detección. Pero en ese caso los neutrinos deberían cambiar de tipo a lo largo de los 148 millones de kilómetros que separan el Sol de la Tierra, y no a lo largo de los 30 metros de este experimento. Los nuevos hallazgos, sin embargo, quizás iluminen la falta de neutrinos muónicos en los

rayos cósmicos que llegan a la Tierra, misterio al que se le da el nombre de problema "atmosférico" de los neutrinos. A principios de agosto se reemprendió de nuevo el experimento del DNLC.

MADHUSREE MUKERJEE

Agujero negro

En la galaxia M87

Los agujeros negros son objetos cósmicos que han sufrido un colapso gravitatorio y cuya masa, millones o incluso miles de millones de veces la solar, ocupa un espacio no mayor que nuestro sistema solar. Su campo gravitacional es tan poderoso, que ni la materia ni, siquiera, la luz que caigan en ellos podrán volver al universo exterior.

Durante años los astrónomos han buscado ansiosamente signos de su existencia. El *Telescopio Espacial Hubble* ha proporcionado la señal más vigorosa hasta la fecha de su realidad. Un equipo de astrónomos dirigido por Holland Ford y Richard Harms llevó a cabo las observaciones.

Estaban estudiando con el *Hubble* las regiones interiores de la M87, una enorme galaxia elíptica situada en el Cúmulo de Virgo a unos 50 millones de años luz de la Tierra. Allí encontraron por casualidad un disco de gas, hasta entonces desconocido, que, a 60 años luz del centro de la galaxia, gira a una velocidad de 550 kilómetros por segundo, casi veinte veces la velocidad a la que la Tierra da vueltas alrededor del Sol.

A partir de este movimiento extraordinariamente veloz, han calculado que el gas gira alrededor de una masa central que posee entre dos y tres mil millones de masas solares. El disco está orientado perpendicularmente a los extensos chorros de gas que dispara el centro de la M87, exactamente tal y como la teoría astrofísica predice.

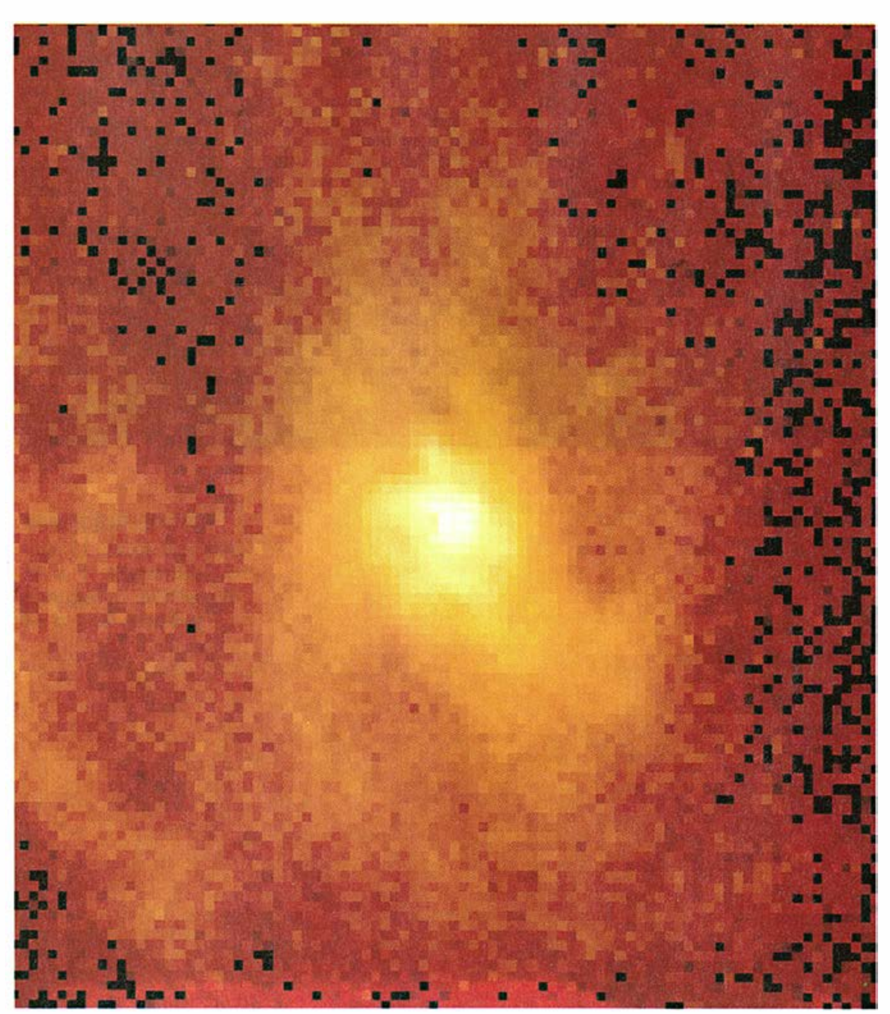
Las nuevas observaciones llegan treinta años después de que Edwin E. Salpeter y Yakov B. Zeldovich propusieran que la materia que cayese en los agujeros negros podría aportar la energías de los cuásares y las radiogalaxias. Cuando los astrónomos dieron en sospechar que los cuásares sólo representan un período extremadamente activo del desarrollo inicial de muchas galaxias, cayeron en la cuenta de que en los núcleos de la mayoría de las galaxias grandes tendrían que permanecer agujeros negros latentes.

El largo chorro de gas radioemisor que emana del centro de M87 señalaba a ésta como un lugar donde era particularmente probable que se hallase un agujero negro de gran masa. A finales de los años setenta, los estudios del movimiento estelar en el núcleo de M87 indicaron que podría haber una aglomeración de estrellas alrededor de ese objeto. Las imágenes que tomó Tod Lauer con el *Hubble* antes de su reparación óptica reforzaron la suposición. Pero el disco de gas encontrado por Ford y Harms ofrece un argumento mucho más convincente. En vez de tener que medir el movimiento de las estrellas cerca del agujero —procedimiento embarullado y poco concluyente—, abordaron una medición mucho más simple, la de la rotación de lo que parece ser un solo disco que gira.

Pero la búsqueda de agujeros negros sigue basándose fundamentalmente en pistas indirectas. Ni siquiera el reparado *Hubble* puede resolver el agujero negro propiamente dicho, cuyo radio mide unos cinco mil millones de kilómetros, 1/100.000 del tamaño de la parte del disco que ve el *Hubble*. Pequeñez y movimiento rápido del disco que descartan todo objeto que no sea un agujero negro.

Harms planea una sesión complementaria con el *Hubble* para determinar las velocidades de zonas más profundas del disco, lo que debería proporcionar un argumento casi a prueba de bomba a favor del agujero negro. Podrán luego los astrónomos ponderar si los aparentemente exóticos y monstruosos agujeros negros no son en realidad sino un resultado bastante corriente de la formación de galaxias. Martin Rees, por ejemplo, arguye que los agujeros negros de gran masa se desarrollaron rutinariamente durante el proceso en el que unas vastas nubes de gas se fundieron y formaron las galaxias en los primeros tiempos del universo, hace miles de millones de años.

COREY S. POWELL



En el corazón de la galaxia M87 hay un disco en rápida rotación. Lo descubrió el Telescopio Espacial Hubble. Es probable que el gas caliente gire alrededor de un agujero negro no visto que se encuentre en el centro

Fotos y texto: Kōji Nakamura

Con antenas por alfombra

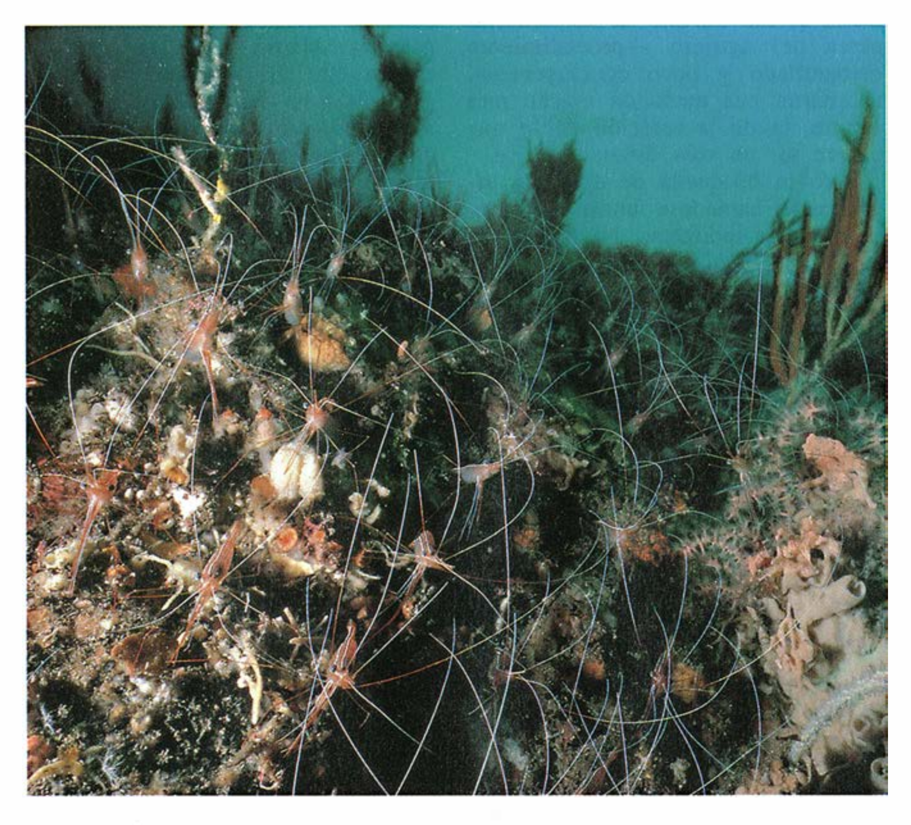




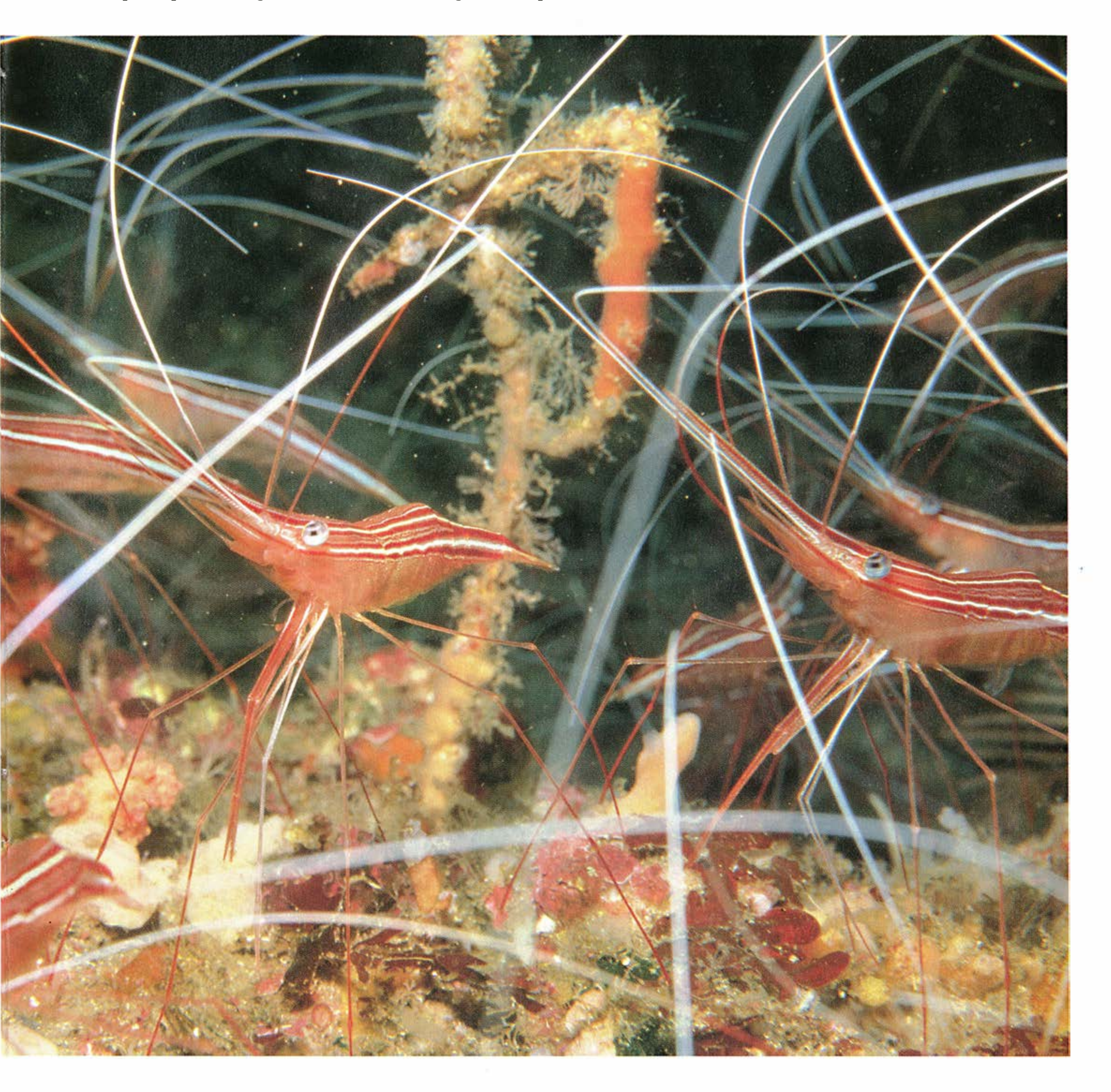
FOTO SUPERIOR
distancia focal: 15 mm
diafragma: F = 16
exposición: 1 segundo
película: ISO = 64

FOTO DE LA DERECHA
distancia focal: 55 mm
diafragma: F = 8
exposición: 1/80 segundos
película: ISO = 25

Entre los ecosistemas del bentos marino hay uno reservado para exquisitos: las cuevas sumergidas. La primera vez que me adentré en ese medio apareció ante mis ojos un universo de colores planos. Encanto que, sin embargo, desaparece conforme penetramos en el interior, de una fauna escalonadamente pobre y menguante. Pero es un mundo, aunque inhóspito, sobrecogedor, donde el tiempo ha perdido su movimiento y la soledad no aquieta. Allí buscan refugio los criptófilos, animales adaptados a las oquedades, intersticios y anfractuosidades.

Fueron los españoles los primeros que llamaron la atención sobre la vida en las cuevas subterráneas, a principios de siglo. También la oceanografía ha progresado gracias a su investigación sobre las comunidades de cuevas submarinas de las islas mediterráneas. Pero aquí, en Japón, tenemos el mar de Jogasaki, y en el mar una isla llamada Honshu, o Principal, donde convergen la corriente Negra y la corriente de las Kuriles. Hay, en el subsuelo, cuevas y entrantes por los que merodean bancos de crustáceos.

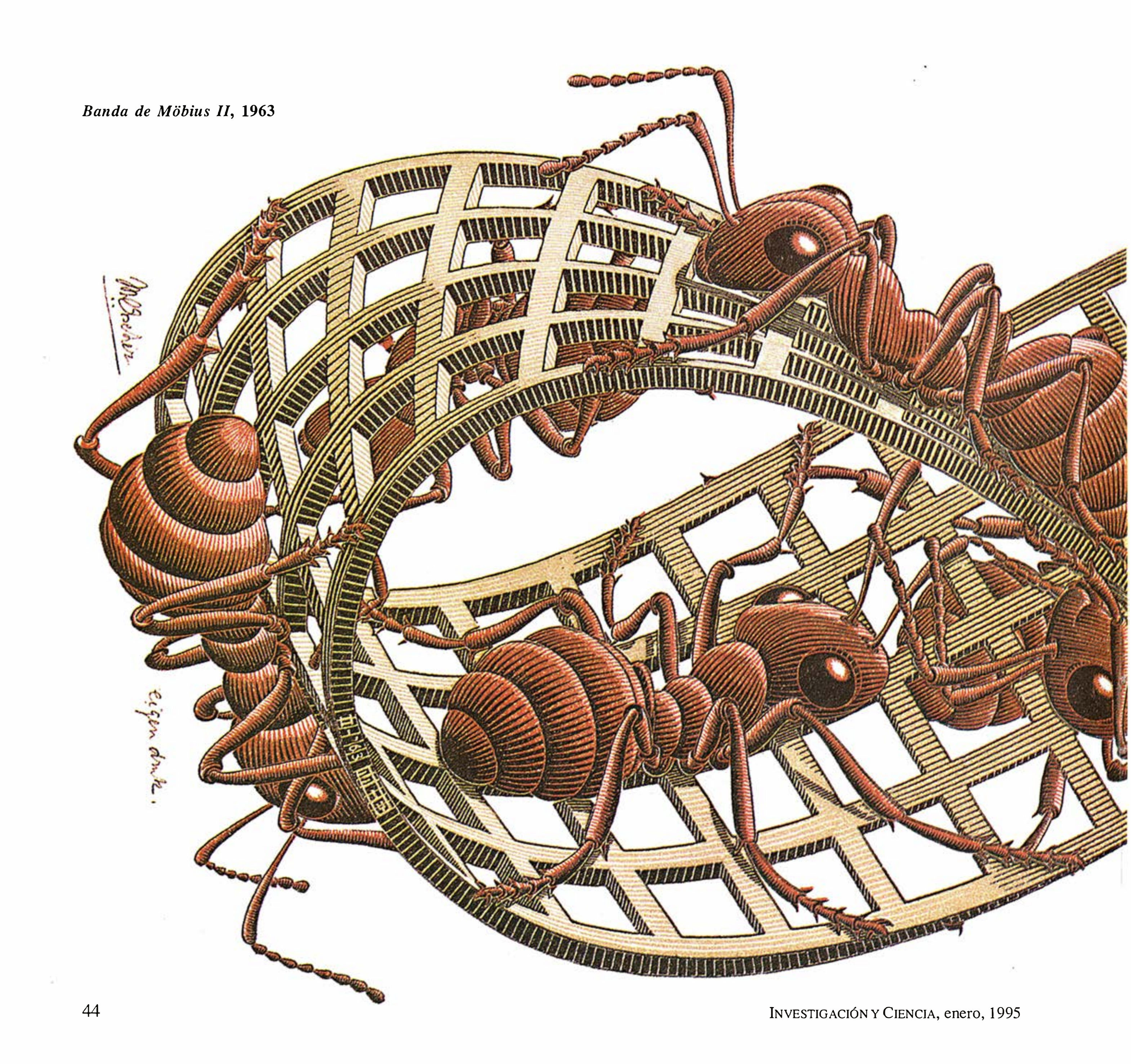
No fue fácil fotografiar el tapiz que tejían las gambitas, que huían de la luz del foco en un ordenado desconcierto de antenas. Insatisfecho de mis resultados, volví un mes más tarde. Habían desaparecido. Momentáneamente, me imagino. No se entendería que abandonaran por siempre esas aguas ricas en nutrientes.



Las metáforas de Escher

Los grabados y los dibujos de M. C. Escher dan expresión a nociones abstractas de las matemáticas y las ciencias

Doris Schattschneider



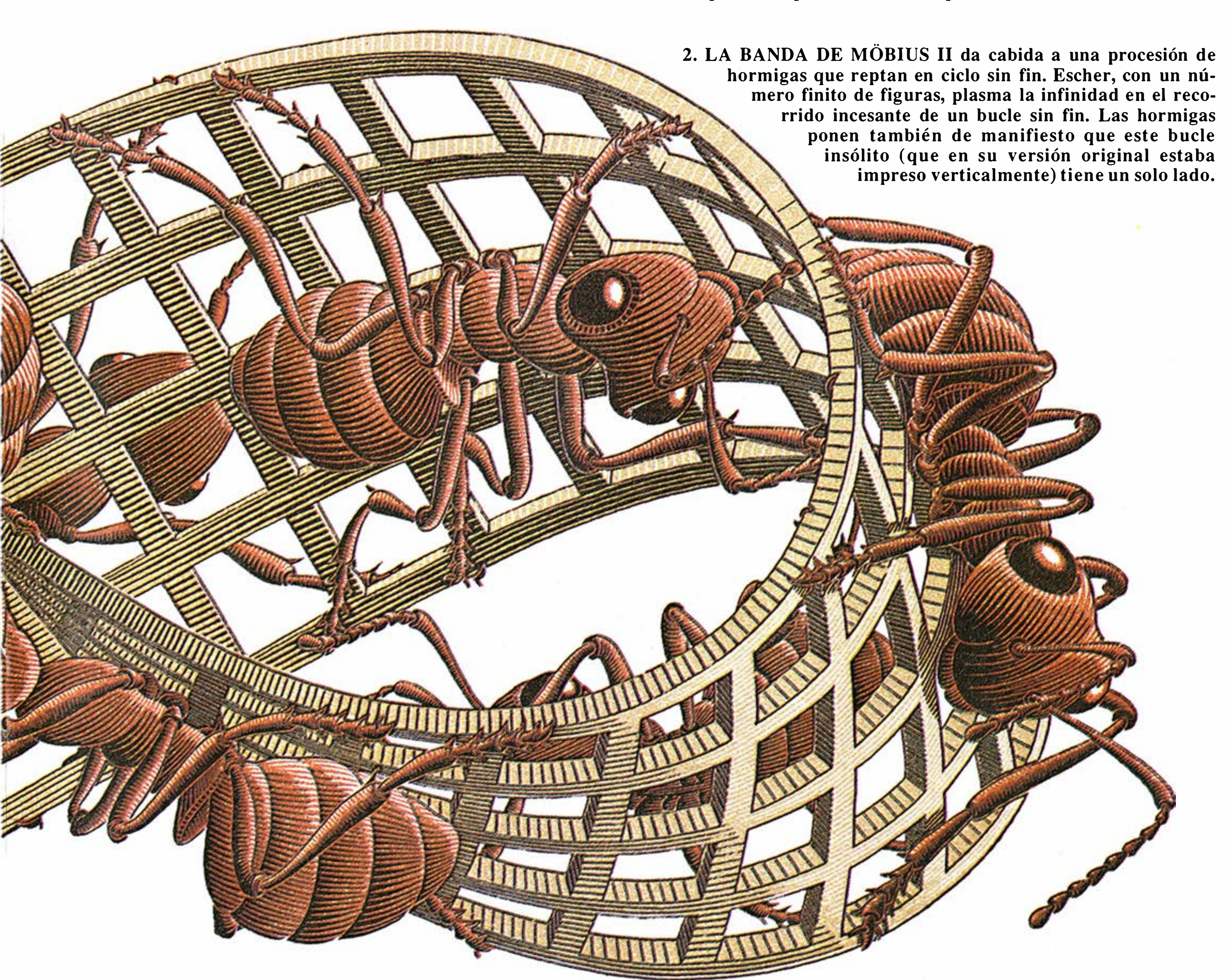
***Banda de Möbius II*, 1963**

Maurits Cornelis Escher solía hablar de su incapacidad para entender las matemáticas, confesándose "absolutamente lego en ciencias exactas". Mas, ya de niño, sintióse intrigado por el orden y la simetría. Esta fascinación le llevaría más tarde a estudiar los motivos de las losetas de la Alhambra de Granada, a observar los dibujos geométricos de los artículos de matemáticas y a investigar ideas originales, exclusivamente suyas, sobre las pavimentaciones de un plano.

La atención que prestó a la coloración de sus dibujos de teselas entrelazadas se anticipó a trabajos posteriores de matemáticos y cristalógrafos en el campo de la simetría de colores. La exposición de sus obras, coincidente con el Congreso Internacional de Matemáticas de 1954, y la publicación de su primer libro *The Graphic Work of M. C. Escher* en 1959 hizo vibrar en matemáticos y científicos una cuerda que todavía resuena con vigor. Escher escribió que una de las principales fuerzas impulsoras de su trabajo era un "vivo interés por las leyes geométricas contenidas en la naturaleza que nos rodea". Sus ideas, expresadas en obras gráficas, han proporcionado metáforas visuales impresionantes para ideas fundamentales de las ciencias. Nació en 1898 en la ciudad holandesa de Leeuwarden. Era el hijo menor de un ingeniero civil y se crió en Arnhem, con cuatro hermanos. Tres de ellos siguieron carreras científicas o técnicas, pero Escher no era buen estudiante de matemáticas. Animado por el profesor de dibujo de su instituto, se interesó por las artes gráficas, empezando por hacer estampas en linóleo.

Autorretrato, **1943**

1. M. C. ESCHER ante el espejo, en este dibujo con tinta litográfica.



2. LA BANDA DE MÖBIUS II da cabida a una procesión de hormigas que reptan en ciclo sin fin. Escher, con un número finito de figuras, plasma la infinidad en el recorrido incesante de un bucle sin fin. Las hormigas ponen también de manifiesto que este bucle insólito (que en su versión original estaba impreso verticalmente) tiene un solo lado.

En 1919 ingresó en la Escuela de Arquitectura y Artes Decorativas de Haarlem, con la intención de estudiar arquitectura. Pero cuando le mostró su trabajo a Samuel Jessurun de Mesquita, que enseñaba artes gráficas en dicho centro, fue invitado a centrarse en ese campo. De Mesquita ejerció una profunda influencia sobre Escher, primero como maestro (en las técnicas de tallado) y luego como amigo y colega.

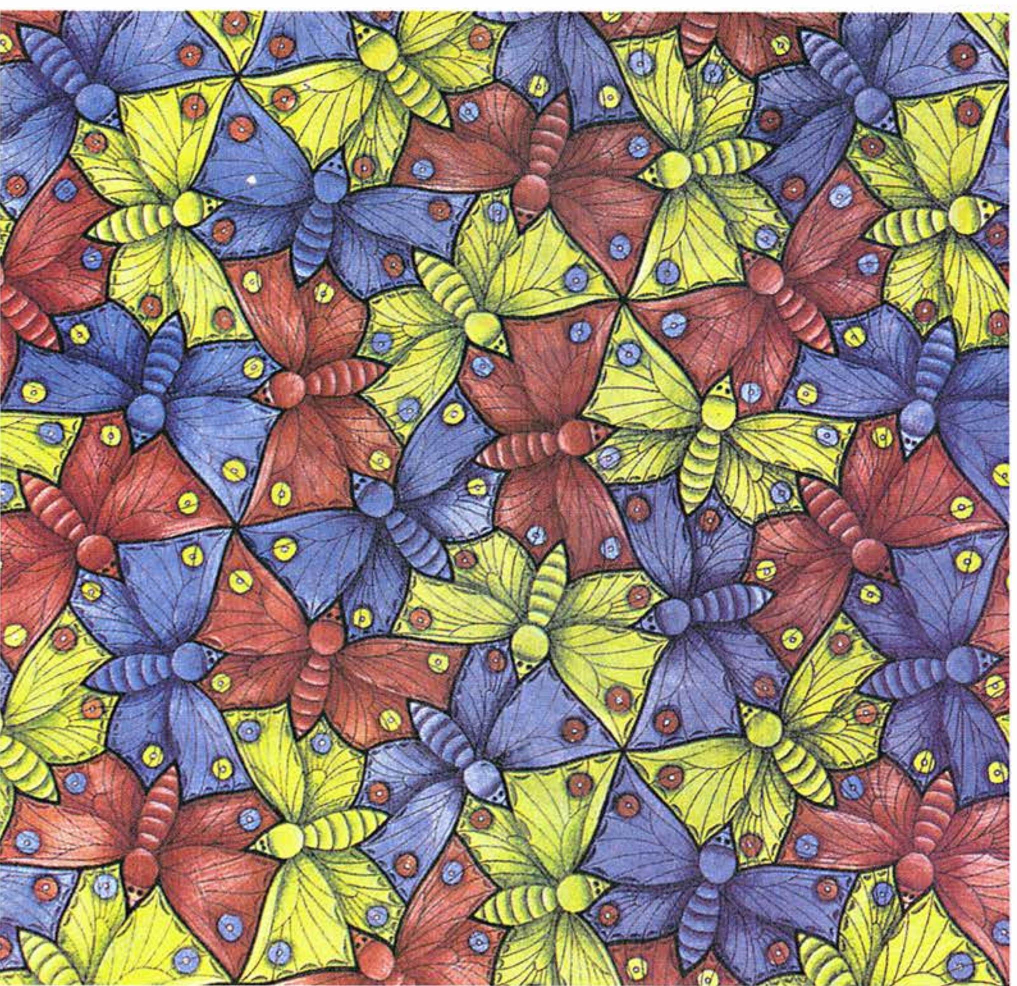
Terminados sus estudios en Haarlem, se estableció en Roma desde donde efectuó numerosas excursiones para tomar apuntes, sobre todo por el sur de Italia. Sus ojos discernían en lo ordinario efectos visuales de gran impacto —detalles arquitectónicos de edificios monumentales vistos desde insólitas perspectivas—, las luces y las sombras proyectadas por escalinatas de pueblecitos, racimos de casas adheridas a las laderas de montes que se hundían en valles lontanos, y en el otro extremo de la escala, detalles diminutos de la naturaleza vistos como a través de una potente lupa. En el estudio de su casa, los bocetos se transformaban en tallas y litografías.

En 1935, la situación política se hizo insoportable. Escher abandonó Italia para siempre en unión de su esposa y sus hijos pequeños. Tras dos años en Suiza, a los que siguieron otros dos en Uccle, cerca de Bruselas, la familia se estableció permanentemente en Baarn, Holanda. Estos años trajeron también un giro abrupto en el trabajo de Escher. A partir de este momento, su inspiración no iba a basarse en lo que observaban sus ojos, sino en lo que veía su imaginación. Buscó dar expresión visual a conceptos y a plasmar las ambigüedades de la observación y el entendimiento humanos. Al hacerlo, se encontró no pocas veces en un mundo regido por las matemáticas.

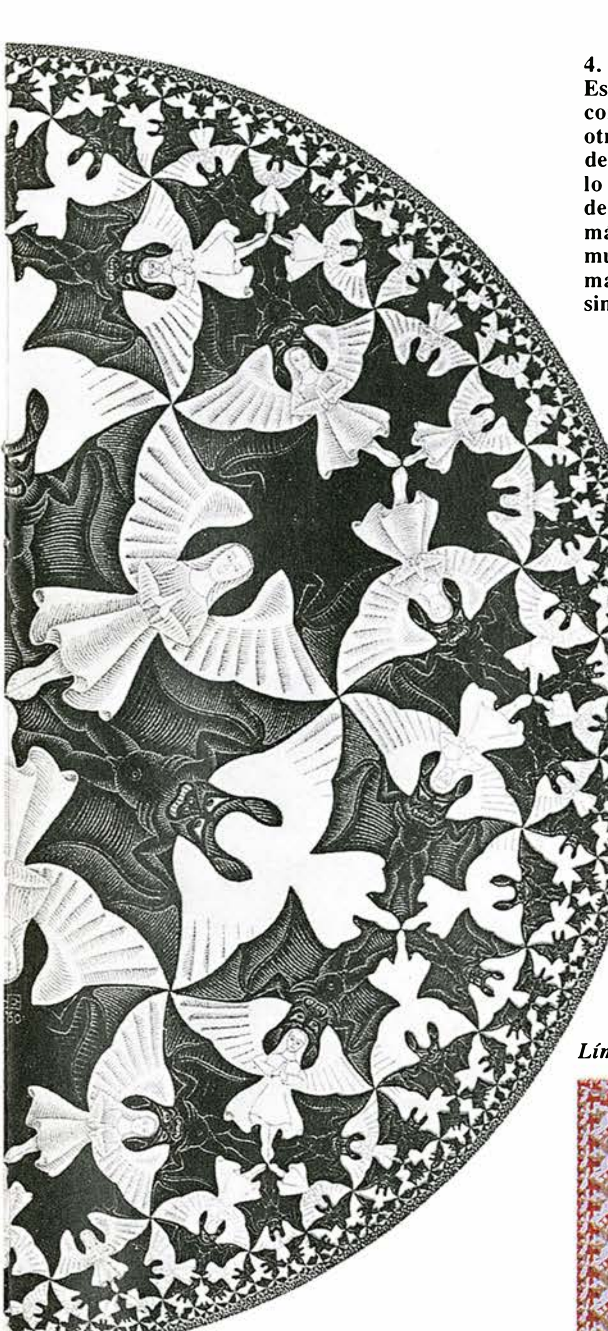
Estuvo fascinado, obsesionado casi, por la idea de "división regular del plano". A lo largo de su vida produjo más de 150 dibujos en color que dieron testimonio de su ingenio y de su originalidad en la creación de figuras que reptaban, nadaban y volaban, llenando el plano con sus clónicos. Estos dibujos son ilustración de simetrías de muchos tipos diferentes. Pero la división del plano era también una forma de captar el infinito. Aunque un teselado como el que se vale de mariposas [*figura 3*] puede en principio proseguir indefinidamente, dando así una sugerencia de la infini-

***Límite circular IV*, 1960**

***Sistema triangular I B_3 Tipo 2*, 1948**



3. LA SIMETRIA es una noción estructural que conforma a muchos modelos físicos y matemáticos. En el dibujo de Escher, las mariposas parecen llenar la página al azar; sin embargo cada una de ellas está situada con gran precisión y rodeada exactamente de igual modo. Siempre son seis (*en colores alternantes*) las que revolotean alrededor del punto donde se tocan las extremidades de las alas delanteras izquierdas; siempre son tres (*de diferentes colores*) las que giran en torno a un punto donde se tocan las alas posteriores derechas. Además de simetría de rotación, el dibujo posee simetría de traslación, basada en un retículo triangular. El motivo puede repetirse ilimitadamente en todas las direcciones, proporcionando así una metáfora del infinito. La atención que prestaba al coloreado fue precursora de descubrimientos matemáticos en simetría de color.

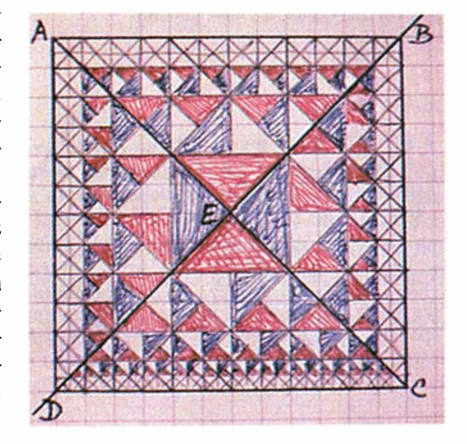


4. LA DUALIDAD es tal vez el tema que más prevalece en los últimos grabados de Escher. En matemáticas, a cada enunciado se le puede asociar su negación, y a cada conjunto, su complementario; en cada caso, el objeto y su dual definen el uno al otro. En *Límite circular IV* no hay perfiles; los contornos de los ángeles y de los demonios se definen mutuamente. Cada uno es, a la vez, figura y fondo (Escher nos lo recuerda omitiendo detalle en la mitad de las figuras). Para nuestros ojos euclídeos, las figuras de esta pavimentación hiperbólica parecen distorsionarse más y más al disminuir de tamaño. Sin embargo, medidas por la geometría intrínseca al mundo del grabado, cada ángel tiene exactamente la misma forma y el mismo tamaño, y otro tanto vale para cada diablo. Un número infinito de copias se repite, sin salirse de los confines del círculo.

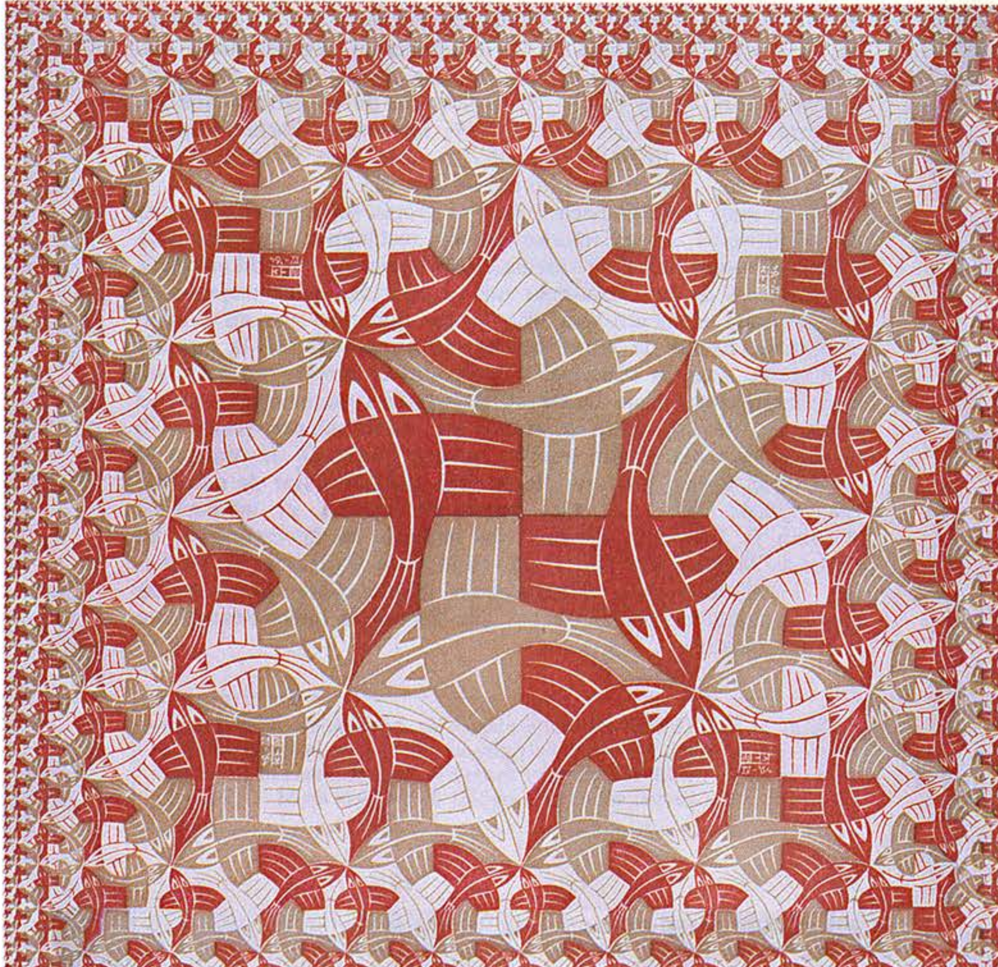
tud, Escher se dispuso a encerrar el infinito en los confines de una sola página.

"Quien se sumerge en el infinito, tanto en tiempo como en espacio, adentrándose más y más sin detenerse, necesita, al pasar raudo, puntos fijos, hitos, pues de lo contrario su movimiento resulta indistinguible de la quietud", escribió. "Es preciso que gradúe su universo en unidades de cierta longitud, en compartimentos que se repitan uno tras otro en interminable sucesión."

Tras terminar varios grabados en los cuales las figuras van sin fin disminuyendo de tamaño conforme se aproximan a un punto de fuga central [*figura 9*], buscó un procedimiento para plasmar en dirección contraria la reducción progresiva. Quería figuras que se repitieran eternamente, aproximándose siempre a un contorno que las circundara, pero sin llegar a alcanzarlo jamás. En 1957, H. S. M. Coxeter le envió a Escher una separata de un artículo publicado en una revista matemática, en el cual Coxeter ilustraba la simetría planar con algunos dibujos de Escher. Este descubrió en él una figura que le

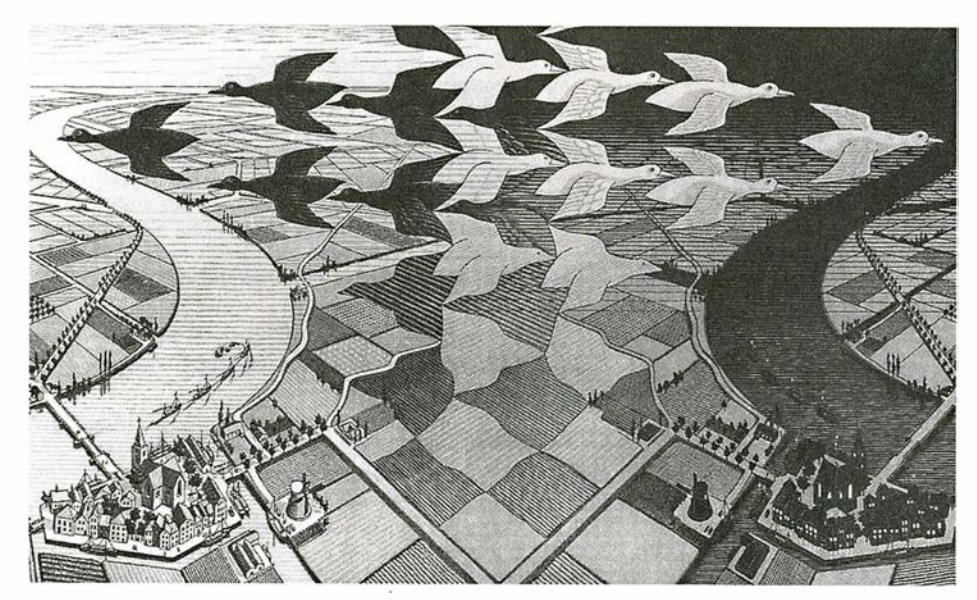


***Límite cuadrado*, 1964**



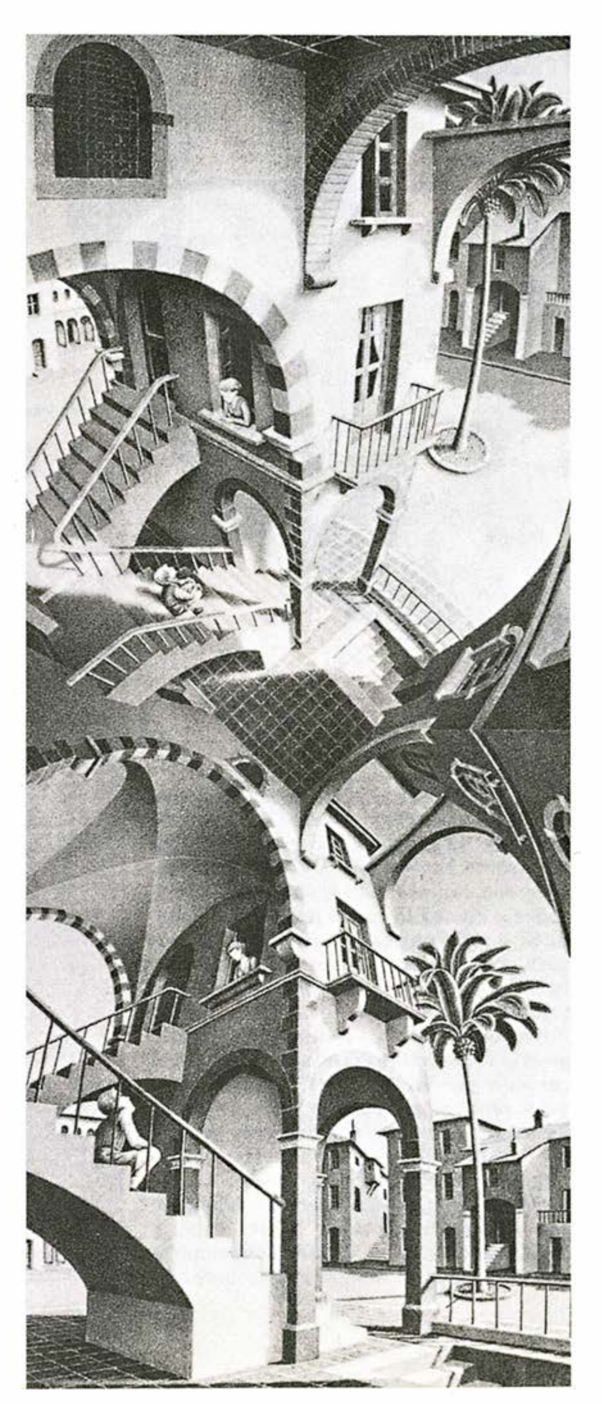
5. LA SIBISEMEJANZA queda ilustrada en el grabado *Límite cuadrado*, construido mediante un esquema recursivo invención del propio Escher. Los algoritmos recursivos son conjuntos de instrucciones aplicadas a un objeto con el propósito de producir nuevos objetos, y luego a los nuevos para producir otros, y así sucesivamente. El producto final es sibisemejante si todos los objetos finales son iguales al original, salvo por cambios de escala, orientación o posición. Un boceto (*arriba a la derecha*) enviado por Escher a H. S. M. Coxeter para explicar el grabado pone de manifiesto que la retícula subyacente se vale de una división recursiva de triángulos isósceles. Al ejecutar el grabado, Escher vació del bloque de madera un triángulo con vértice en el centro del cuadrado y cuya base es uno de los lados del cuadrado e imprimió el bloque cuatro veces.

Día y Noche, 1938



6. LA DIMENSION es esa noción que separa punto, recta, plano y espacio. Para poner de manifiesto las ambigüedades en la percepción de dimensión, Escher sacó partido de la página impresa, que siempre ha de engañar al observador cuando ofrezca la imagen de una escena tridimensional. En *Día y Noche*, el escaqueado plano de los campos de labor al pie del grabado se metamorfosea en dos bandadas de gansos. El grabado ilustra también la noción de cambio topológico, en el cual las figuras son deformadas sin ser cortadas ni perforadas. También se aprecian la simetría por reflexión y la dualidad o complementariedad: los gansos negros vuelan sobre una villa iluminada por el sol, mientras que los blancos aletean sobre la imagen nocturna y simétrica de la misma escena.

Alto y Bajo, 1947



7. LA RELATIVIDAD afirma que lo que ve un observador está influido por el contexto y por el punto de observación. En la litografía *Alto y Bajo*, Escher presenta dos vistas de la misma escena. En la mitad inferior, el observador se encuentra en el patio; en la mitad superior, el observador mira hacia abajo. Alejémonos ahora del grabado. El cuadrilátero pavimentado que ocupa el centro, ¿corresponde a suelo o a techo? Escher lo utiliza de ambas formas, a fin de casar las dos vistas. Es imposible ver con lógica la figura completa. La escena también pone de manifiesto que, al pegar visiones parciales para formar un todo global, pueden aparecer contradicciones.

produjo "toda una conmoción": una teselación hiperbólica de triángulos, que plasmaba exactamente el efecto que tanto buscaba. Tras un estudio cuidadoso del diagrama, Escher pudo discernir las reglas de teselación, en las cuales arcos circulares forman ángulo recto con la circunferencia de un círculo que los encierra. A lo largo de los tres años siguientes, Escher produjo cuatro grabados distintos basados en este tipo de retícula, el último de los cuales fue *Límite circular IV*.

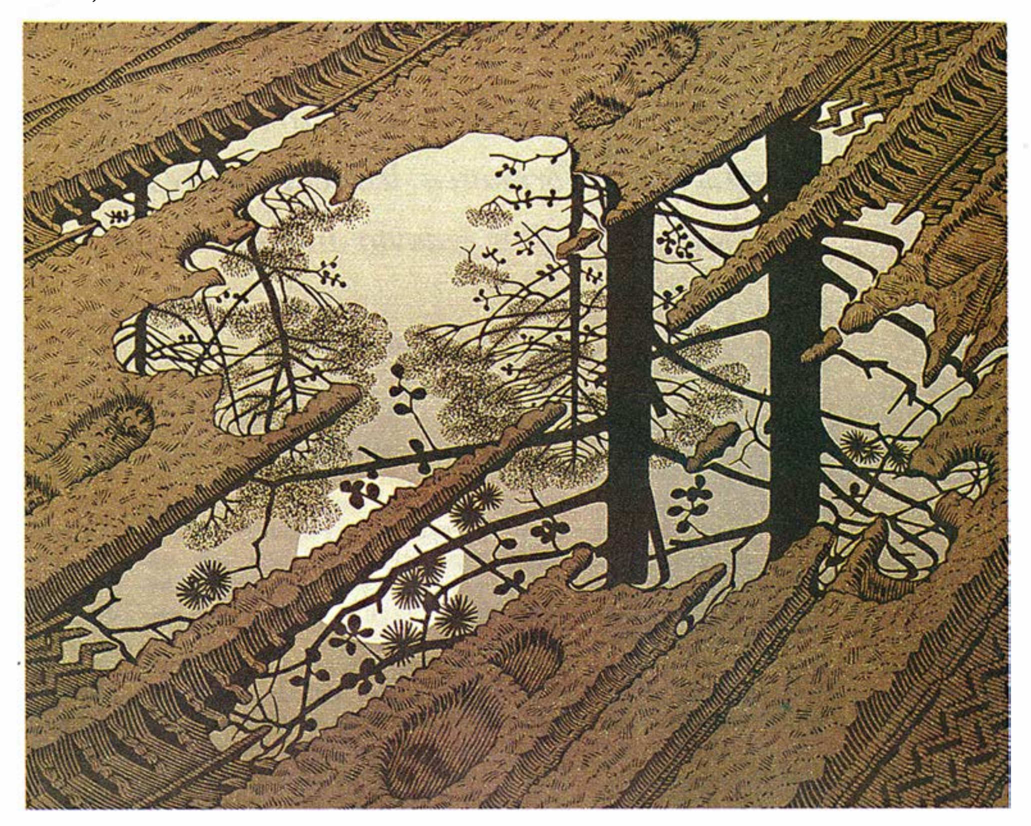
Cuatro años después Escher ideó su propia solución al problema de la infinitud en el interior de un rectángulo. Su algoritmo recursivo —una serie de instrucciones aplicadas repetidamente a un objeto— da por resultado una regularidad sibisemejante en la cual cada elemento está relacionado con otro por un cambio de escala. Escher le envió a Coxeter un esbozo de la retícula subyacente, con la disculpa: "Temo que el motivo no resultará muy interesante, visto desde su perspectiva matemática, porque realmente es tan sencillo como un recubrimiento plano. No obstante, la tarea de dar con un método adecuado de realizar el tema de la forma más sencilla posible produce dolor de cabeza." Hace pocos años, en unos cursos de verano, el matemático William P. Thurston, director del Instituto de Investigación en Ciencias Matemáticas de la Universidad de California en Berkeley, ilustró la noción de teselación sibisemejante precisamente con un reticulado así, sin saber que Escher lo había descubierto con anterioridad.

DORIS SCHATTSCHNEIDER, profesora de matemáticas en el Colegio Moravian de Bethlehem, Pennsylvania, se doctoró en la Universidad de Yale en 1966. Su libro *Visions of Symmetry* fue culminación de un proyecto de investigación sobre los estudios de simetría de M. C. Escher.

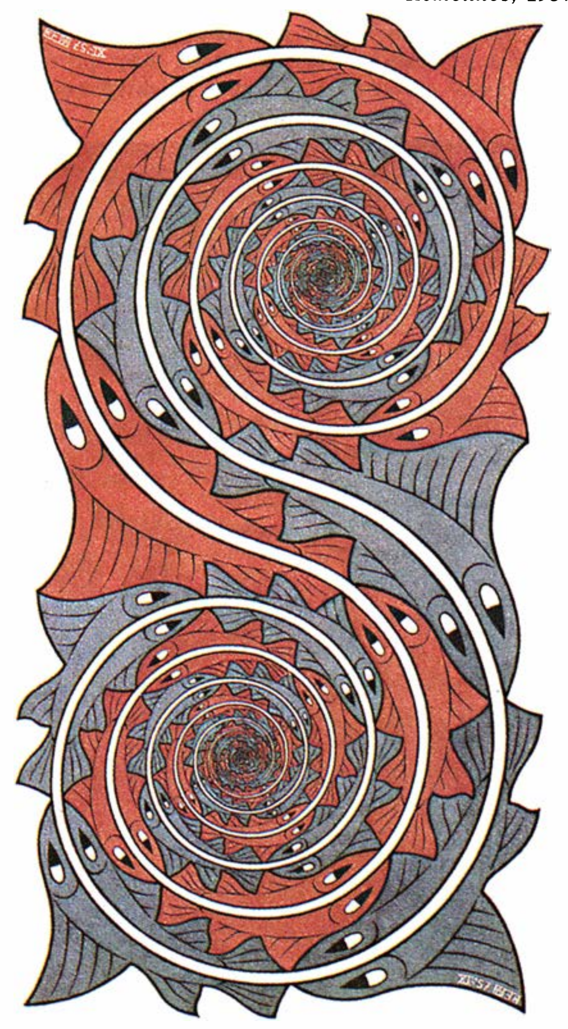
8. LA REFLEXION permite observar fenómenos demasiado pequeños, demasiado alejados o demasiado oscuros para la percepción directa. *Charco* dirige nuestra mirada hacia una senda del bosque hollada por botas y neumáticos; aun así, en el charco se nos muestra también el perfil de los árboles, arqueados hacia lo alto en un cielo iluminado por la Luna. Escher nos recuerda los mundos invisibles situados por debajo, por detrás y por encima de nuestra limitada visión.

9. EL INFINITO es confinado en el espacio finito de un grabado en *Remolinos*. El artista ha dibujado una proyección plana de la curva (llamada loxodrómica) que se va trazando sobre el Globo por una senda que corte a todos los meridianos con el mismo ángulo. Como saben todos los navegantes, al seguir esta "línea de rumbos" resulta una espiral interminable, que se ciñe cada vez más en torno a los polos.

***Charco*, 1952**



***Remolinos*, 1957**



Los motivos sibisemejantes proporcionan ejemplos de figuras que poseen dimensión fraccionaria, o fractal, ambigüedad que sin duda hubiera causado gozo a Escher. En 1965 éste confesaba que: "No puedo evitar burlarme de todas nuestras inflexibles certidumbres. Resulta muy divertido, por ejemplo, confundir deliberadamente las dimensiones dos y tres, el plano y el espacio, e inyectar humor en la gravedad." Escher era un maestro en la confusión de dimensiones, como en *Día y Noche*, en la cual los campos bidimensionales de las granjas se metamorfosean misteriosamente en ánades tridimensionales. Le encantaba poner de manifiesto las ambigüedades y contradicciones inherentes a una práctica habitual en las ciencias: la de pegar entre sí varias vistas locales de un objeto para formar un todo global [*figura 7*].

Hacia el final de su vida (murió en 1972) escribió: "Por encima de todo, me complazco en el contacto y la amistad de los matemáticos a que todo ello ha dado lugar. Con frecuencia me han proporcionado nuevas ideas y a veces ha existido incluso una interacción entre nosotros."

BIBLIOGRAFIA COMPLEMENTARIA

THE GRAPHIC WORK OF M. C. ESCHER. M. C. Escher. Ballantine Books, 1971.

THE MAGIC MIRROR OF M. C. ESCHER. B. Ernst. Random House, 1976.

ANGELS AND DEVILS. H. S. M. Coxeter, en *The Mathematical Gardner*. Dirigido por David A. Klarner. Prindle, Weber and Schmidt, 1981.

M. C. ESCHER: HIS LIFE AND COMPLETE GRAPHIC WORK. Dirigido por J. L. Locher. Harry N. Abrams, 1982.

M. C. ESCHER: ART AND SCIENCE. Dirigido por H. S. M. Coxeter, M. Emmer, R. Penrose y M. L. Teuber. North-Holland, 1986.

ESCHER ON ESCHER: EXPLORING THE INFINITE. M. C. Escher. Traducción al inglés por Karin Ford. Harry N. Abrams, 1989.

VISIONS ON SYMMETRY: NOTEBOOKS, PERIODIC DRAWINGS, AND RELATED WORK OF M. C. ESCHER, D. Schattschneider. W. H. Freeman and Co., 1990.

Redes informáticas seguras

Aunque las redes de ordenadores no pueden presumir de su seguridad informática, los realizadores del sistema Atenea no han visto fallar todavía ninguno de sus protocolos

Jeffrey I. Schiller

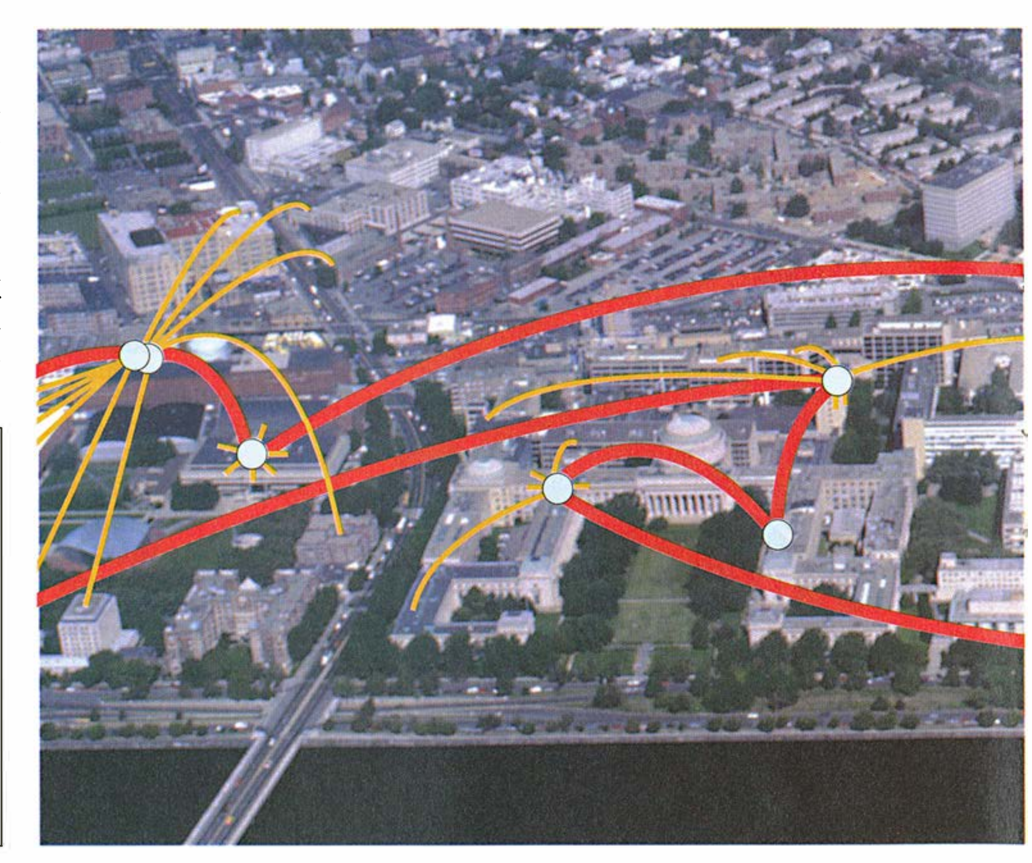
La ampliación a escala planetaria de la red Internet ha traído consigo cambios notables en las formas de comunicación y de hacer negocios. Mas, parejamente a esta expansión de la aldea global, ha ido creciendo el número de individuos de honradez menos que dudosa. Y al diseñar gran parte de los equipos y programas que constituyen Internet sólo se pensó en usuarios dignos de confianza. El resultado es que la privacidad de las comunicaciones a través de la red sufre la invasión de furtivos, que pueden leer o alterar la información almacenada.

Los análisis y debates públicos recientes sobre seguridad informática se han centrado, sobre todo, en atacantes activos, que introducen en la red datos falsos para perturbar sus funciones normales. Muchos fisgones y vándalos informáticos se han abierto paso hasta ordenadores conectados a Internet aprovechando ciertos puntos vulnerables del sistema que, pese a ser bien conocidos, no se han reparado todavía. En la actualidad están adquiriendo importancia cada vez mayor los ataques pasivos, que se limitan a escuchar el tráfico en la red sin perturbar su flujo. A decir verdad, una de las amenazas más insidiosas que pesan sobre la seguridad de las redes modernas es la facilidad con que pueden "pincharse" para fisgar en ellas. Así, mientras el acceso a los ordenadores dependa de contraseñas transmitidas por una red, un atacante pasivo podrá explotar una mina de oro digital sin más que escuchar a hurtadillas.

Los administradores de los sistemas informáticos del Instituto de Tecnología de Massachusetts (MIT) han estado habiéndoselas con problemas de seguridad desde hace más de 10 años. En 1983, en colaboración con IBM y con Digital Equipment Corporation, el MIT emprendió el proyecto Atenea, para poner a disposición de alumnos y docentes medios informáticos avanzados. Atenea se funda en el modelo "cliente-proveedor": las estaciones de trabajo de los usuarios (los clientes) efectúan casi todos los cómputos, mientras que máquinas especializadas enlazadas a una red que cubre todo el recinto universitario, llamada MITnet, se ocupan del almacenamiento de archivos, de su impresión y de otras tareas. (De entonces acá, este modelo se ha convertido en paradigma ampliamente utilizado en buena parte de la Internet.)

Durante el diseño del sistema Atenea, tuvimos muy presente que la red iba a entrañar un gran problema de seguridad. Hablando sin tapujos, dimos por supuesto que nuestros estudiantes tardarían muy poco en averiguar la forma de programar las

JEFFREY I. SCHILLER ha gestionado la red informática del Instituto de Tecnología de Massachusetts (MIT), de la cual forma parte el sistema Atenea de computación distribuida, concebida en 1984. Cursó ingeniería eléctrica en el MIT. Coautor del sistema de acreditación Cerbero, dirige los análisis sobre seguridad del Internet Engineering Steering Group y la Internet Engineering Task Force. Schiller, que ha participado en el desarrollo de las normas de Internet para reforzar la privacidad del correo electrónico, acaba de preparar una versión del popular programa de encriptación PGP.



estaciones de trabajo de la universidad para hacerlas actuar como eficaces detectoras y registradoras de contraseñas. Para amortiguar ese riesgo, desarrollamos el sistema de certificación Cerbero, que evita definitivamente tener que enviar contraseñas legibles a través de puntos vulnerables de la red. Este tocayo del can tricéfalo que otrora guardaba la entrada de los infiernos es el corazón de la seguridad de Atenea. Sin embargo, a diferencia de su antecesor, nuestro perro guardián todavía no ha fallado una sola vez en su vigilancia.

Los problemas de seguridad que plantean las redes informáticas tienen una doble vertiente, conceptual y práctica. Resulta que unos pocos tramos de hilo de cobre, de línea coaxial o de fibra óptica pueden crear confusión en las categorías mismas utilizadas para decidir qué es lo seguro y qué no lo es. Pensemos en la noción de fronteras de un sistema informático. Los especialistas han venido clasificando sus preocupaciones en dos categorías: la seguridad perimétrica, que impide a los situados en el "exterior" acceder al "interior" de un sistema, y la seguridad interna, que impide que los usarios interiores se obstruyan mutuamente o infrinjan las normas de seguridad del sistema.

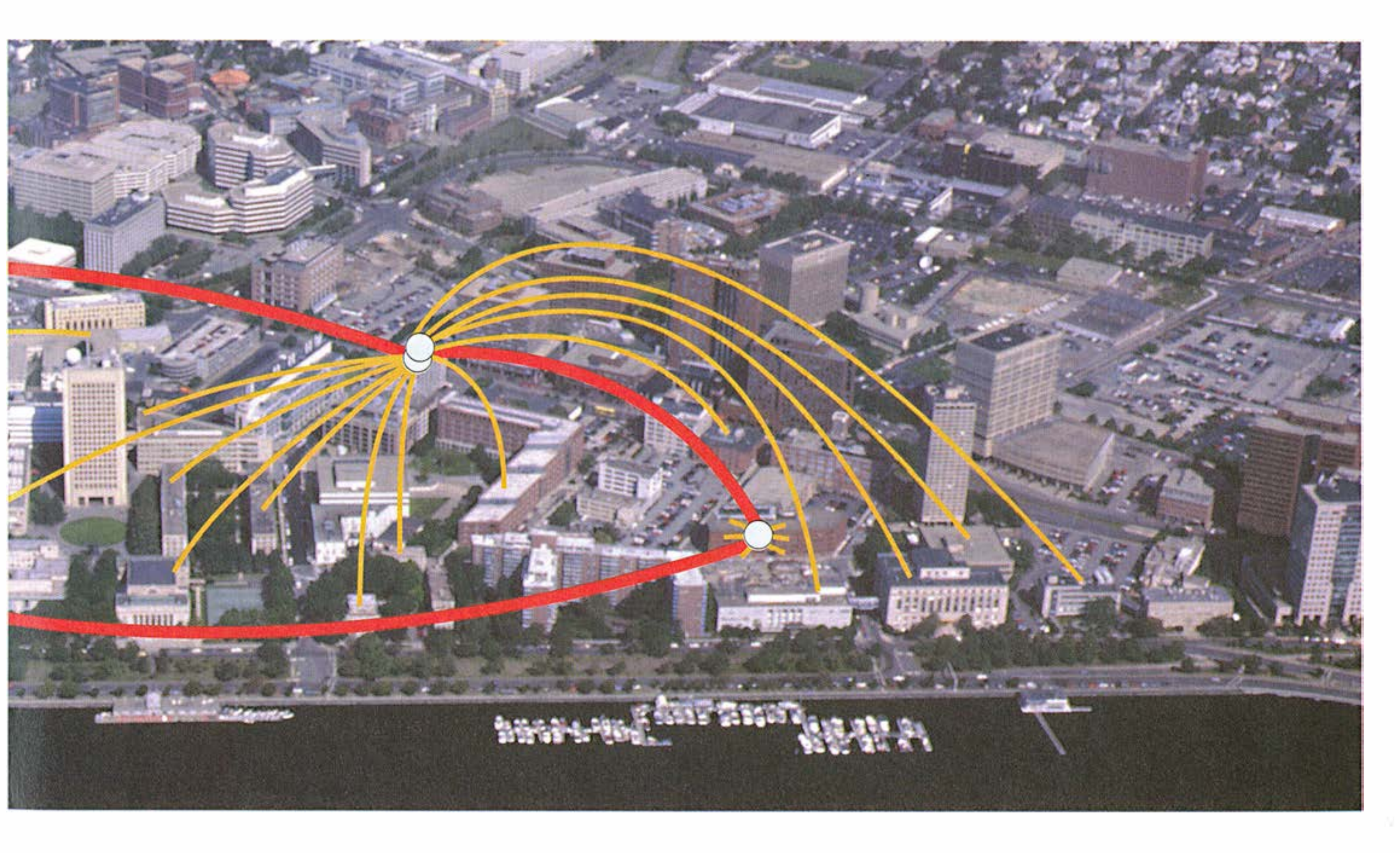
Resulta bastante sencillo definir el perímetro de un sistema informático tradicional, basado en un gran ordenador. La unidad central de proceso, la memoria y las unidades de disco están "dentro"; todo lo demás se halla fuera. Los dispositivos de entrada/salida, como los lectores de tarjetas, los lectores de cintas y los terminales constituyen el perímetro; toda información que se introduzca en el sistema ha de pasar a su través. Por ejemplo, para que los usuarios puedan utilizar un terminal han de acreditarse tecleando un número de cuenta y una contraseña. Unicamente el titular de una cuenta concreta debe conocer su contraseña, por lo cual basta la conjunción de ambos datos para identificar al usuario ante el sistema.

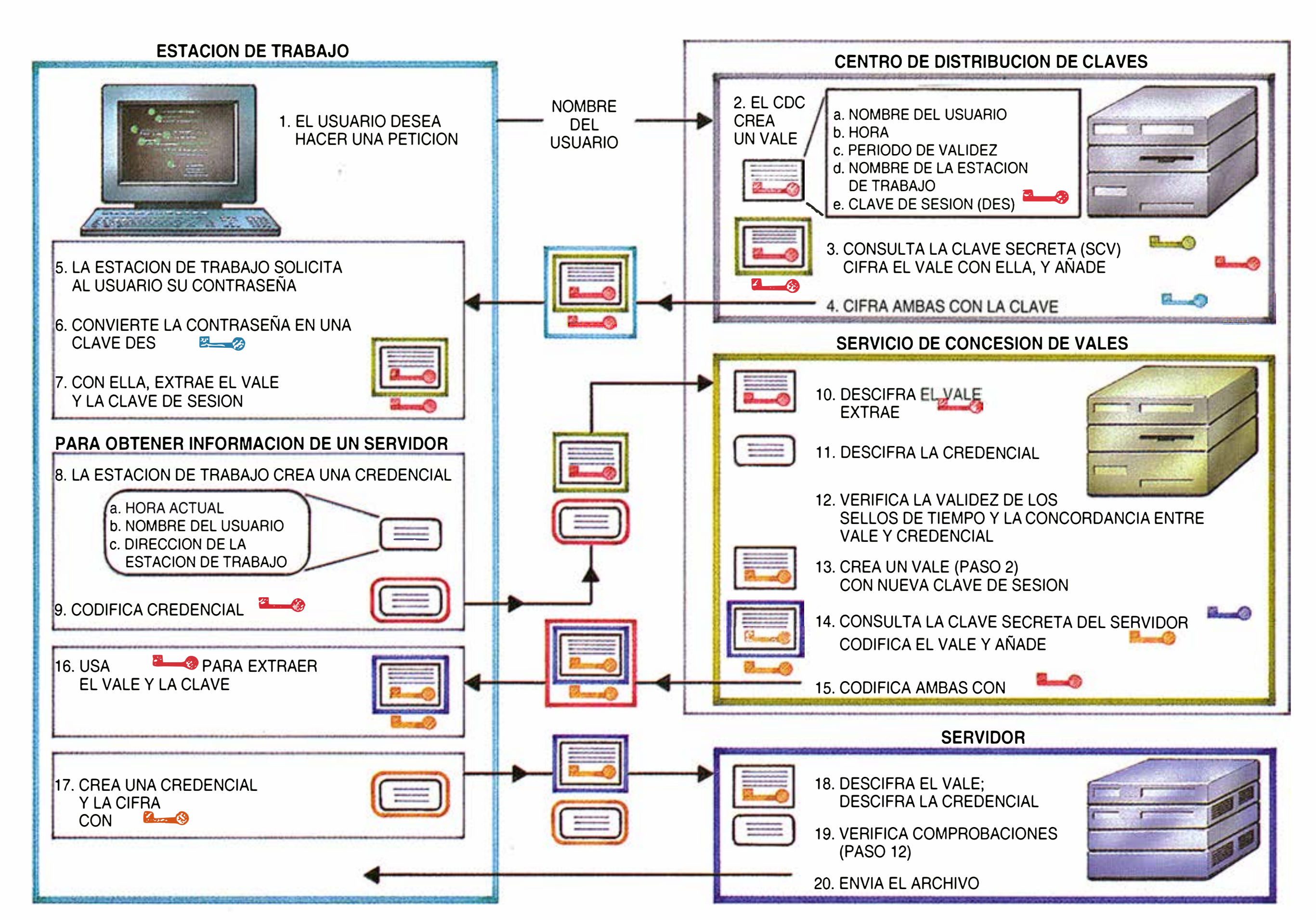
Importa señalar que la acreditación del sistema ante el usuario es implícita: el usuario da por supuesto que se está comunicando con el ordenador deseado. Si el terminal está directamente conectado a un ordenador específico, el usuario sabe que el terminal "habla" en nombre de ese sistema. El usuario que se comunica telefónicamente con un ordenador central, marcando un número en su módem, confía en que la compañía telefónica conecte su llamada con el sistema informático correspondiente al número marcado. (Los usuarios y los gestores de un sistema confían también en el mecanismo de distribución del número telefónico adecuado.)

En tanto no se subvierta la seguridad interna del ordenador, la confianza mutua entre usuario y ordenador central se mantiene hasta que el usuario se desconecta. El sistema da por supuesto que todas las pulsaciones de teclas que ha recibido proceden del usuario. Este, por su parte, supone análogamente que toda la información aparecida en su pantalla procede del sistema informático adecuado.

Estos axiomas en que se funda la seguridad de los grandes ordenadores fallan tan pronto entran en juego redes y estaciones de trabajo. En pri-

1. RED INFORMATICA del recinto universitario del Instituto de Tecnología de Massachusetts: interconecta más de 10.000 ordenadores. Su columna vertebral consiste en un anillo de fibra óptica con capacidad para 100 megabits por segundo (*parcialmente mostrado en líneas rojas*). Encaminadores (*círculos azul celeste)* situados de forma estratégica por el recinto universitario transfieren datos, a medida que se les va requiriendo, entre más de 100 redes Ethernet de área local (*vemos algunas, en amarillo*) y la columna vertebral de fibra óptica. Alrededor de 1200 de los ordenadores acoplados a MITnet forman parte del entorno informático Atenea; las interacciones entre ellos están protegidas mediante protocolos criptográficos.





2. LOS PROTOCOLOS CERBERO se basan en claves de cifrado, que sólo conocen quienes deban participar en una transacción, para proteger la información enviada a través de una red abierta. El diagrama de la figura muestra la secuencia que se desarrolla cuando los usuarios solicitan acceso al sistema ("log in") para utilizar los archivos propios. Tan sólo se precisa la contraseña del usuario en el primer intercambio; las peticiones subsiguientes se basan en la clave de sesión (*rojo*) que es compartida por el usuario y por el servicio de concesión de vales.

mer lugar, ya no resulta fácil definir el perímetro de un sistema conectado en red. En lugar de un sistema informático individual, a cuyo alrededor podemos dibujar una línea, un entorno de computación distribuida consta de muchos sistemas independientes conectados mediante enlaces cuya fiabilidad o seguridad es incierta.

La mayoría de las redes propenden a ser grandes y difíciles de tornar seguras. En el MIT, los cables se cruzan y entrecruzan por todo el campo universitario, extendiéndose incluso hasta las residencias de estudiantes. Pura y simplemente, resulta demasiado caro encerrar todas las líneas en armarios telefónicos y en conductos invulnerables. Además, casi todas las redes se basan en un sistema de difusión generalizada ("broadcast"): cada uno de los ordenadores conectados a un cable de la red tiene acceso a toda la información que fluye por ese cable. Los soportes físicos y lógicos instalados en los ordenadores se encargan de clasificar los datos que van pasando; por lo general, cada uno de ellos procesa tan sólo la información que le está dirigida. Sin embargo, un intruso que tenga control de un ordenador puede fácilmente programarlo para que reciba, ora todos los datos del tramo de red al que se halla conectado, ora todos los datos dirigidos a algún otro ordenador. Tal intruso podría también enviar información y hacer, a la vez, que parezca provenir de algún otro lugar.

La facilidad con que un intruso puede llevar a cabo tales actos ilícitos revela que la red no se encuentra, estrictamente hablando, dentro del perímetro de seguridad del entorno de computación distribuida. Si la red se halla en el exterior de tal perímetro, será preciso proteger los paquetes de datos que transportan información entre las estaciones de trabajo y los servidores. Cada paquete ha de verificarse y certificarse en el momento de atravesar el perímetro de seguridad representado por la red.

En el sistema Atenea, esta certificación está encomendada a Cerbero, un sistema de programación distribuida que se vale de una serie de intercambios de información cifrada antes de permitir al usuario el acceso a los servidores. Cerbero proporciona, asimismo, una serie de verificaciones criptográficas para garantizar que los datos transferidos entre estaciones y servidores no están corrompidos, sea por accidente o por acción intencionada de terceros.

La versión de Cerbero utilizada en el MIT se vale del éstandar de encriptación de datos (Data Encryption Standard, DES) para cifrar sus comunicaciones. El DES opera fragmentando los mensajes en bloques discretos de información (compuestos, normalmente, de ocho caracteres, o sea, de 64 bits) que transforma en bloques de texto cifrado de acuerdo con una "clave" de 56 bits. La descodificación exige utilizar la misma clave para reconvertir los bloques de texto cifrado en el mensaje original. Antes de enviar un paquete, la estación de trabajo puede cifrarlo mediante una clave conocida solamente por esa estación y por el servidor al

cual están dirigidos los datos. La interceptación de las transmisiones de la red de nada le servirá a un atacante, porque, a falta de la clave, cada paquete es puro galimatías digital.

Además, la naturaleza del algoritmo DES hace sencilla la detección de cualquier alteración hostil de la información que pasa por la red. Cualesquiera cambios que sufra un paquete provocarán que su descodificación produzca "basura", vale decir, caracteres aleatorios sin ninguna relación con el mensaje original. Los mensajes corruptos de este tipo son fácilmente detectables, y la estación de trabajo o el servidor pueden desecharlos sin más y solicitar una retransmisión.

Los métodos criptográficos para enviar información de forma segura a través de redes no seguras constituyen tan sólo los cimientos de Cerbero. Hemos edificado sobre ellos una superestructura de protocolos capaces de identificar a los individuos que solicitan servicios informáticos. Puede hacer también que los servidores y las estaciones de trabajo posean una clave secreta conocida exclusivamente por ambas máquinas y por nadie más (de modo que las protecciones criptográficas sean realmente eficaces). Los protocolos empiezan con una clave secreta DES de 56 bits para cada usuario. Cada servicio de la red tiene también una clave secreta. A los humanos no nos resulta cómodo recordar una secuencia aleatoria de 56 bits (que corresponde a un número de 20 dígitos); en consecuencia, el sistema permite elegir una contraseña de seis a 128 caracteres. Una etapa de codificación adicional convierte la contraseña en una clave DES. Por ejemplo, la ristra de caracteres "mydog" se traduce en 12, 322, 343, 883, 628, 311, 502 (10101011000000001110000100111100 110111111110111110111001111001110).

Las claves de todos los usuarios y servidores son conocidas tan sólo por un servidor especial, que recibe el nombre de central de distribución de claves (CDC); interviene éste en todas las transacciones por medio de certificados digitales cifrados, a los que nosotros llamamos "vales". Cuando los usuarios se sientan ante una estación de trabajo y escriben los nombres de su "login" para darse de alta (el "login" es el nombre por el cual los conoce el sistema), dan comienzo a un intercambio de mensajes entre la estación y el CDC, que acaba por producir un vale para utilizar un determinado servicio de la red. La estación empieza primero por decirle al CDC que el usuario desea efectuar una petición —por ejemplo, utilizar datos almacenados en un servidor de archivos. El CDC crea un paquete de datos —el vale—, que contiene el nombre del usuario, la hora, el intervalo de tiempo en que el vale será válido, el nombre de la estación de trabajo y una clave DES generada aleatoriamente, llamada clave de sesión. Consulta luego la clave secreta correspondiente al servidor de archivos y cifra el vale, de suerte que sólo pueda leerlo el servidor.

A continuación, el CDC codifica este vale ya cifrado con la clave secreta del usuario y remite el resultado a la estación de trabajo. Cuando la estación recibe esta información, le solicita al usuario una contraseña. Convierte la contraseña en una clave DES y la utiliza para descodificar el vale y la clave de sesión que lo acompaña. Tan sólo en el caso de que el usuario le suministre a la estación de trabajo la contraseña correcta, podrá ésta descodificar tal información.

Cuando el usuario desea solicitar información del servidor de archivos, la estación de trabajo le entrega al servidor el vale y un paquete de datos adicional, llamado "credencial". (La credencial está formada por la hora actual, el nombre del usuario y la dirección de la estación de trabajo, todo ello encriptado con la clave de sesión.) El servidor de archivos descodifica el vale y extrae la clave de sesión, con la cual puede entonces descifrar la credencial. Se cerciora de que el usuario y la estación de trabajo mencionados en el vale concuerden con los de la credencial; verifica, asimismo, que las marcas de hora sean válidas. Si todas estas credenciales superan la inspección, el servidor de archivos acepta el vale y procesa la solicitud del usuario.

A excepción del nombre del usuario, toda la información contenida en estas transacciones atraviesa la red en forma cifrada y está, por consiguiente, a salvo de los ojos de fisgones. Por otra parte, además de contribuir a la acreditación del usuario, la clave de sesión incluida en el vale permite al usuario cifrar los datos que pasan entre su estación y el servidor de archivos.

Para situaciones en las que pueda resultar delicado determinar si ha existido corrupción de datos, Cerbero tiene prevista un comprobación de autenticidad de mensajes (CAM). Esta salvaguardia adicional consta de un número de 128 bits, derivado de los datos originales no cifrados, y de una clave secreta, compartida por la

La red Atenea

El proyecto Atenea utiliza dos tipos básicos de elementos informáticos: estaciones de trabajo y servidores. Las estaciones de trabajo albergan las unidades de proceso con potencia suficiente para efectuar los cómputos del usuario, pero carecen de medios para el almacenamiento a largo plazo de la información particular de cada usuario. Los discos rígidos de las estaciones de trabajo locales contienen solamente el sistema mínimo de programas necesarios para conectarlas a MITnet, red informática del recinto universitario. Una vez conectada, la estación de trabajo puede obtener toda una batería de programas de sistema operativo de un servidor de archivos.

Del almacenamiento de archivos, al igual que de otras particularidades del entorno informático, se encargan ordenadores "servidores". Estos ordenadores pueden ser estaciones de trabajo avanzadas, miniordenadores, grandes ordenadores o cualesquiera otras máquinas que los gestores del sistema estimen adecuadas, en la medida en que puedan comunicarse usando los protocolos idóneos. Atenea se vale de servidores para almacenar ficheros de datos de particulares, para hacer funcionar las impresoras, para encargarse del sistema de correo electrónico y para gestionar otros servicios subsidiarios.

Las estaciones de trabajo son muy caras. Por ello, el Instituto de Tecnología de Massachusetts no exige a sus alumnos la adquisición de equipos propios. Los estudiantes tienen máquinas a su disposición en determinados lugares públicos del recinto universitario. Uno de estos puntos de acceso puede albergar entre 12 y 120 estaciones de trabajo. Recientemente, se ha ampliado la MITnet hasta las residencias estudiantiles, interiores o exteriores al perímetro del MIT. Los estudiantes pueden así conectar a la red sus propios ordenadores. Estos sistemas pueden funcionar de forma independiente o bien actuar como estaciones de trabajo integradas en el entorno Atenea.



3. SERVIDORES del entorno Atenea. Residen, en su mayor parte, en este sótano del centro de operaciones de la red informática del MIT. Dentro de la zona de servidores hay una sala cerrada bajo llave ("la mazmorra"), donde está instalado el ordenador encargado del centro de distribución de claves de Cerbero.

ción, cualquier intruso podría recabar de ésta vales y claves de sesión. Estos vales, sin embargo, sólo son utilizables desde dicha estación (porque contienen el nombre de la estación) y cada uno es válido tan sólo durante unas cuantas horas (un máximo de 10 en la configuración del MIT).

El uso de Cerbero y del sistema de cifrado a él asociado no garantiza la seguridad de un entorno de computación distribuida. Son necesarias otras medidas para impedir que los atacantes merodeen por las garantías que Cerbero puede proporcionar. Ante todo, los servidores han de ser seguros por sí mismos. Cerbero proporciona un medio para que las estaciones de trabajo (o los servidores) confirmen las identidades individuales. Sin embargo, después de eso, los servidores han de seguir tomando la decisión adecuada acerca de quiénes han de ser autorizados para acceder a un determinado recurso. Felizmente, la tarea no resulta demasiado difícil. La mayoría de los servidores pueden ser configurados de manera que sólo los gestores del sistema puedan entrar directamente en ellos (esto es, sin solicitar servicios a través del mecanismo de vales). Cuesta menos garantizar la seguridad porque el acceso a ellos queda reservado a personas de confianza.

Y lo que reviste mayor importancia: el propio CDC de Cerbero tiene que ser físicamente seguro. Dispone de una copia de la clave de cada uno de los usuarios, por lo que un intruso que logre acceder al CDC y lea esta información lo que consigue, en esencia, es apoderarse de la contraseña de todo el mundo. De hecho, los intrusos no necesitan modificar ninguna información del CDC propiamente dicho para lograrlo; les bastaría leerlo (¡o leer las cintas de respaldo!). La versión de CDC utilizada en el MIT tiene en su programación ciertas características que le hacen difícil a un intruso la lectura directa de las claves DES; ello no obstante, sigue siendo vital el mantenimiento de la seguridad física, lo que se consigue con vigilancia del CDC y de las cajas fuertes donde se guardan las copias de respaldo, grabadas en cintas.

El CDC tiene que ser seguro, pero debe también operar a todas horas, con el fin de que los usuarios pue-

estación de trabajo y el servidor; la salvaguardia se añade como apéndice a un mensaje. Cuando el mensaje llega, vuelve a computarse el CAM. Si los dos valores coinciden, el destinatario puede estar seguro de que no ha existido manipulación. Si los datos que es preciso enviar desde el cliente al servidor no son confidenciales, una comprobación de autenticidad del mensaje puede bastar para protegerlo de manipulaciones no autorizadas, sin tener que molestarse en cifrar el mensaje entero.

Aunque el protocolo básico de Cerbero es seguro, no se adapta bien a usuarios normales. En funcionamiento, revela una debilidad que pudiera crear dificultades. Cada servicio nuevo requiere un vale diferente, por lo que el usuario podría tener que suministrar una contraseña cierto número de veces durante una sesión. La contraseña corre peligro si se almacena en una estación de trabajo. Son muchas las causas que podrían inducir a un usuario a dejar una estación de trabajo sin darse de baja en ella (y borrando así la contraseña depositada). Un intruso sagaz podría sencillamente acercarse a esa estación desocupada y robar en ella la contraseña.

El problema de la estación desatendida quedaría resuelto solicitando al usuario una contraseña cada vez que haga falta; pero, a la larga, la reiteración de peticiones es igualmente peligrosa. No siempre se puede predecir cuándo necesitará el cliente un nuevo servicio. Por ejemplo, el servidor de la Estafeta Postal de Atenea, que contiene el correo electrónico por leer, difiere del servidor donde residen los archivos ordinarios. En consecuencia, las transacciones con la Estafeta precisan de vales propios, y por ello el usuario podría tener que suministrar una contraseña cada vez que llegase nuevo correo electrónico. Suponiendo que los usuarios no pierdan los nervios al serles reiteradamente exigido que demuestren su identidad, acabarán habituándose a suministrar una contraseña a cualquier programa que se la solicite. Se convierten entonces en presas fáciles para un intruso que proporcione un programa (por ejemplo, un juego de ordenador) que solicite contraseña, la cual, en vez de ser utilizada para acreditación en Cerbero, se almacena para su posterior recuperación.

La solución a este dilema estriba en el servicio de concesión de vales (SCV). El SCV, instalado en el mismo sistema que el centro de distribución de claves, tiene acceso a su base de datos de usuarios, servicios y claves. Los usuarios le proporcionan una contraseña una sola vez, en el momento de causar alta en el servicio, para traer un vale para el SCV desde el CDC. Las peticiones posteriores de vales para otros servicios van al SCV, que las cifra, no con la contraseña del usuario, sino con la clave de sesión que acompañaba al vale SCV inicial.

Por tanto, la contraseña del usuario no tiene por qué estar almacenada en la estación de trabajo. Sólo reside en la memoria el tiempo justo para permitir a la estación de trabajo el descifrado del vale SCV. Si el usuario dejase desatendida su esta-

dan conectar con Atenea. Tal exigencia plantea una antinomia, porque el método tradicional de aumentar la fiabilidad en un ambiente de computación distribuida consiste en replicar los servicios críticos en múltiples máquinas (ubicadas, por lo común, en diferentes localidades), para disponer de un margen de garantía frente a fallos de energía u otros problemas. Sin embargo, cada CDC ha de contar con protección física. Al aumentar el número de centros de distribución de claves se fortalece la fiabilidad, pero al mismo tiempo se multiplica el riesgo de que el sistema corra peligro.

En teoría, Cerbero parece sólido. Pero, ¿qué tal funciona en la práctica? Empezamos a utilizarlo en el MIT en 1986, y la versión actual (la número 4) se viene empleando en el recinto universitario desde 1987. No sabemos de ninguna situación en la que la seguridad de Atenea haya corrido peligro por culpa de un defecto en los mecanismos de seguridad de Cerbero. Sin embargo, Cerbero no protege a los sistemas de fallos habituales, así la adivinación de contraseñas elegidas con poco seso, el engaño a los gestores del sistema por intrusos astutos o los abusos que los individuos de confianza puedan cometer con sus privilegios. Los fallos de los dos últimos tipos son casi imposibles de extirpar. Pero en 1991 instalamos en nuestro programa de cambio de contraseñas un filtro que eliminaba las contraseñas manifiestamente mal elegidas; el filtro no permitía, por ejemplo, la elección de contraseñas fáciles de adivinar (como el nombre propio o una palabra de uso corriente.)

Otras muchas instituciones han adoptado el sistema Cerbero. Muy notablemente, la Open Software Foundation (organización que promociona sistemas operativos estandarizados) ha hecho de la versión 5 un componente esencial de su entorno de computación distribuida. Esta versión de Cerbero posee algunas características nuevas; además, está menos sujeta a las restricciones del sistema operativo Unix y puede, por consiguiente, adaptarse a otros entornos informáticos. En la actualidad está en período de pruebas.

También se están estudiando y sopesando las posibilidades de Cerbero en la comunidad Internet, donde se demanda a voz en grito la protección de los datos en su viaje alrededor del mundo. Estamos trabajando en nuevos protocolos que pudieran permitirnos construir sistemas distribuidos como Atenea capaces de servir a centenares de miles de usuarios, y no sólo a los 25.000 que ahora estamos atendiendo. Es probable que la criptografía de clave pública (una forma de cifrado que utiliza dos claves, pero sólo exige que una de ellas se mantenga secreta) desempeñe un papel crucial en estos sistemas. Además, las técnicas de clave pública podrían constituir uno de los procedimientos para reducir las exigencias materiales de seguridad de los centros de distribución de claves y de minimizar el riesgo que entraña la multiplicación de copias del CDC.

El éxito logrado por Cerbero hasta la fecha emana de cierto número de fuentes. Una de ellas, evidente, es la experiencia: Atenea fue el primer sistema informático ampliamente distribuido que se ocupó de los problemas de seguridad y acreditación. Otra, su simplicidad. Los usuarios que conectan con Atenea ven los mismos mensajes que si estuvieran comunicándose mediante un sistema multiusuario típico; de hecho, en su mayor parte ignoran por completo el funcionamiento interno de Cerbero. Por último, Cerbero es un sistema abierto: el MIT practica una política de distribución gratuita (que aplica también a otros paquetes de programas, como el sistema X Windows). En consecuencia, tanto los científicos informáticos como los posibles usuarios pueden disponer del código fuente de Cerbero. Muchos de ellos han ofrecido su experiencia y conocimientos para refinarlo y para depurarlo de errores o posibles puntos flacos, y esperamos que continúen haciéndolo así.

BIBLIOGRAFIA COMPLEMENTARIA

USING ENCRYPTION FOR AUTHENTICATION IN LARGE NETWORKS OF COMPUTER. R. M. Needham y M. D. Schroeder en *Communications of the ACM*, vol. 21, n.º 12, páginas 993-999; 1978.

SECTION E.2.1: KERBEROS AUTHENTICATION AND AUTHORIZATION SYSTEM. S. P. Miller, B. C. Neuman, J. I. Schiller and J. H. Saltzer en *Project Athena Technical Plan*. MIT Project Athena, diciembre de 1987.

KERBEROS: AN AUTHENTICATION SERVICE FOR OPEN NETWORK SYSTEMS. J. G. Steiner, B. C. Neuman y J. I. Schiller en *Usenix Conference Proceedings*, Dallas, Texas, págs. 191-202; febrero de 1988.

DISTRIBUTED COMPUTING, IMPLEMENTATION AND MANAGEMENT STRATEGIES. Dirigido por Raman Khanna. Prentice Hall, 1993.

¿Por qué hablan solos los niños?

No debiera reprenderse a los niños cuando hablan solos y en voz alta. Tales soliloquios les ayudan a controlar su comportamiento y a adquirir nuevas habilidades

Laura E. Berk

Lo habrá advertido cualquier padre, maestro, niñera u observador ocasional: los niños pequeños hablan solos; a veces, tanto como hablan con los otros, si no más. A tenor de la situación, esos soliloquios, o habla privada en el argot de los psicólogos, representan entre el 20 y el 60 % de los comentarios que hace un niño menor de 10 años. Padres y educadores suelen ver erróneamente en esa conducta un signo de desobediencia, distracción o inestabilidad mental. Pero la verdad es que el habla privada forma parte esencial del desarrollo cognoscitivo de todos los niños. El reconocimiento de este hecho debiera presidir la enseñanza de niños normales y de los que tienen problemas de aprendizaje.

Aunque los soliloquios infantiles se remontan probablemente a los orígenes del lenguaje, hubo dos factores que impidieron a los psicólogos y pedagogos reconocer su importancia hasta los años sesenta. Me refiero al socialismo ruso de los años treinta y a la autoridad indiscutida de un eximio teórico cognoscitivo occidental. Hace más de seis décadas, Lev S. Vygotsky daba cuenta pública del interés del hablar solo. Pero en aquella época el régimen socialista perseguía a muchos intelectuales, y abundaban las purgas en las universidades e institutos de investigación.

Ese clima de terror ahondó las diferencias entre psicólogos. Hubo quienes tildaron a Vygotsky de renegado. Colegas y discípulos le abandonaron. Según recuerda uno de sus alumnos, el partido programó un "debate" crítico contra las ideas de Vygotsky. Pero en 1934, antes de que pudiera replicar, profundizar en sus trabajos o defender su posición ante el partido, murió de tuberculosis. Dos años después, se proscribía su obra.

LAURA E. BERK enseña psicología en la Universidad estatal de Illinois. Se licenció en la californiana de Berkeley y se doctoró por la de Chicago. Investigadora prolífica, es coeditora de *Private Speech: From Social Interaction to Self-Regulation* y autora de libros de texto universitarios.

El nombre de Vygotsky no atravesó el telón de acero. Psicólogos y pedagogos de Occidente seguían negando, por influencia de Jean Piaget, que el habla privada desempeñara un papel positivo en el desarrollo cognoscitivo normal. En los años veinte, antes de que Vygotsky iniciara sus investigaciones, Piaget había realizado una serie de estudios seminales en los que había registrado cuidadosamente las expresiones verbales de niños de tres a siete años en el Instituto J. J. Rousseau de la Universidad de Ginebra. Además de comentarios sociales, Piaget identificó tres tipos de manifestaciones que no se entendían bien o que no estaban dirigidas claramente a un oyente: los niños repetían, a modo de juego, sílabas o sonidos, mantenían soliloquios y lo que denominó monólogos colectivos.

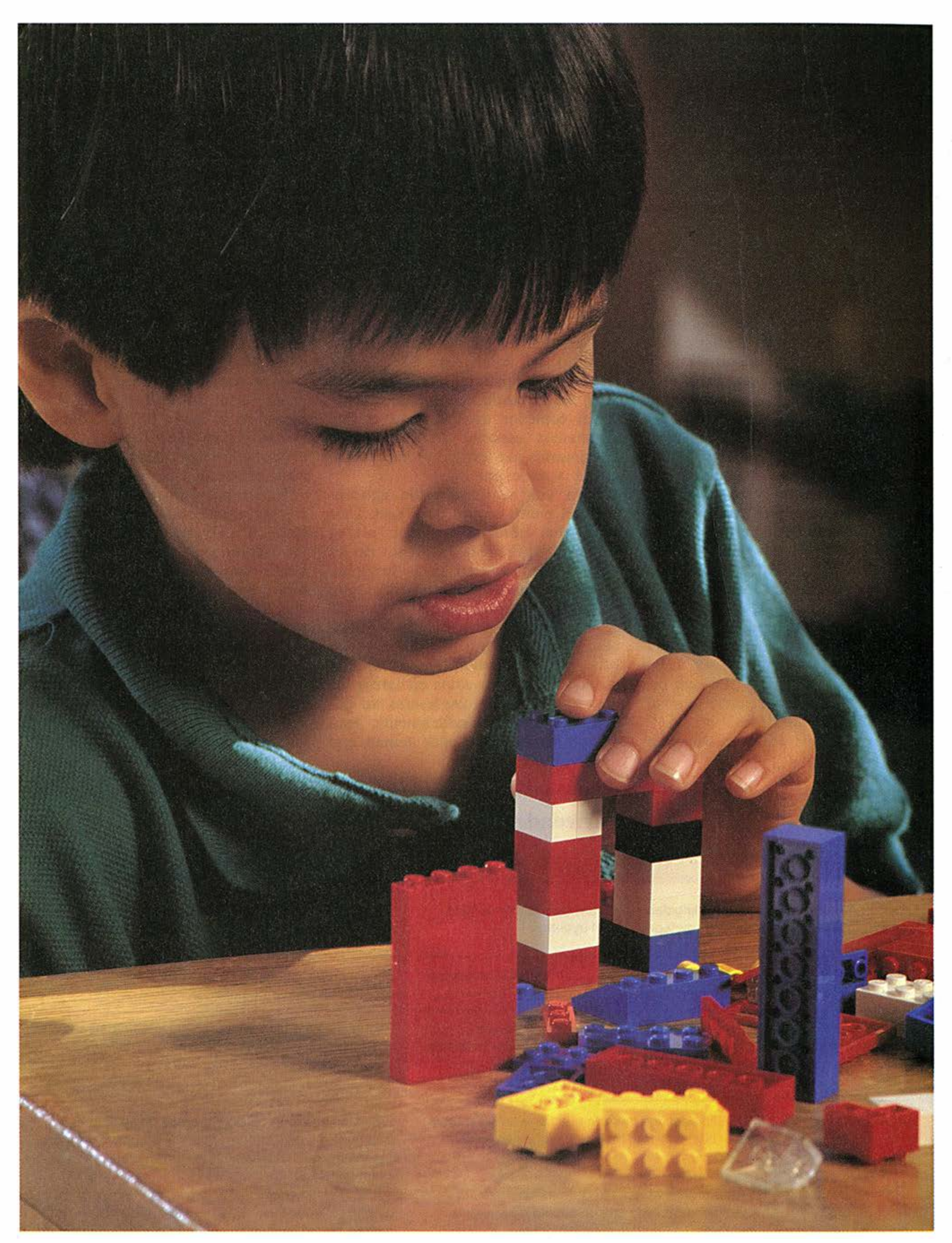
Piaget calificó de egocéntricos a estos tres tipos de habla, pues eran, en su opinión, un producto de mentes inmaduras. Los niños pequeños, razonaba, se empeñan en el habla egocéntrica porque tienen dificultad para imaginar la perspectiva de otro. Su charla es, en buena parte, un soliloquio sin apenas función comunicativa; se trata de mero acompañamiento, complemento o refuerzo de la actividad motriz, cuando no consiste en incongruencias: la expresión verbal de un niño estimula el habla en otro, pero no se espera que el compañero escuche o comprenda. Piaget creía que el habla privada desaparecía gradualmente, a medida que aumentaba la capacidad infantil de una interacción social real.

Pese a las críticas de algunos sectores, las ideas de Piaget imperaron hasta que el trabajo de Vygotsky llegó a Occidente en los años sesenta. Tres años después de la muerte de Joseph Stalin, ocurrida en 1953, Nikita S. Khrushchev censuró el "imperio del terror" y anunció una política de mayor libertad intelectual. Se levantó la prohibición de las obras de Vygotsky. En 1962 apareció en Estados Unidos la traducción al inglés de una compilación de sus ensayos, *Thought and Language*. En menos de un decenio, un equipo dirigido por Lawrence Kohlberg, de la Universidad de Harvard, había recogido pruebas sólidas en apoyo de las ideas de Vygotsky.

A finales de los años setenta, algunos psicólogos comenzaron a distanciarse de Piaget, al tiempo que se traducía una parte más gruesa del corpus vygostkyano. Sumado todo ello a los resultados de Kohlberg, emergió una eclosión de nuevas investigaciones. Desde mediados de los años ochenta se ha triplicado el número de estudios sobre habla privada realizados en Occidente. Casi todos, el mío incluido, corroboran las opiniones de Vygotsky.

El psicólogo ruso describía una estrecha vinculación entre la experiencia social, el habla y el aprendizaje. Afirmaba que los aspectos de la realidad que un niño está preparado para dominar se hallan en la zona de desarrollo proximal (o potencial). Cita un conjunto de tareas que el niño no puede acometer sin la guía de un adulto o de un compañero más hábil; cuando el pequeño comenta con su mentor una tarea a realizar, éste le ofrece directrices y estrategias habladas. El niño incorpora el lenguaje de esos diálogos en su soliloquio y lo usa para encauzar su propio empeño.

"El momento más interesante del curso del desarrollo intelectual", escribió Vygotsky, "... acontece cuando el habla y la actividad práctica, dos líneas de desarrollo hasta entonces completamente independientes, convergen". La dirección del desarrollo no sigue una trayectoria piagetiana, en



1. HABLAR SOLOS supone, para los niños, dirigir su propio comportamiento, adquirir nuevas habilidades y desenvolverse en situaciones desacostumbradas. Cuando un niño se enfrenta a una tarea nueva, recitará en voz alta los aspectos del problema que le resultan enigmáticos. Conforme van ganando en capacidad, el habla privada se convierte en una murmuración inaudible. Por fin, cuando se han ejercitado las operaciones cognoscitivas necesarias para la realización de esa tarea, el niño piensa las palabras en voz baja. La autora ha estudiado ese proceso, siguiendo los trabajos pioneros de Lev S. Vygotsky.

que la comunicación social acaba sustituyendo a las expresiones egocéntricas. Para Vygotsky, la comunicación social precoz precipita el habla privada; la comunicación social da lugar a todos los procesos cognoscitivos superiores, específicos del ser humano. Al relacionarse con miembros maduros de la sociedad, los niños aprenden a dominar las actividades y a pensar de una forma congruente con su cultura.

Conforme va siendo dueño de su conducta, el pequeño no precisa ya la expresión completa del soliloquio; el sujeto, después de todo, es un oyente entendedor. Por consiguiente, los niños omiten las palabras y las frases que aluden a lo que ya conocen sobre una situación dada. Expresan sólo los aspectos que continúan pareciendo enigmáticos. Una vez ejercitados en las operaciones cognoscitivas, empiezan a "pensar palabras" más que a decirlas. Paulatinamente, el habla privada se va convirtiendo en habla interna, silenciosa, en esos diálogos conscientes que mantenemos con nosotros mismos cuando pensamos y actuamos. Pero el impulso de hablar solos nunca desaparece. En cualquier momento de nuestra vida en que nos encontremos ante una situación insólita o frente a actividades absorbentes, resurgirá el habla privada. Es una herramienta que nos ayuda a superar obstáculos y a adquirir nuevas habilidades.

En los últimos años se han diseñado dos programas encaminados a confirmar y completar las ideas de Vygotsky. El de la autora y el dirigido por Rafael M. Díaz, de la Universidad de Stanford. Los dos partimos de cuestiones similares: ¿Hablan solos todos los niños? ¿Les ayuda a guiar sus acciones? ¿Tiene su origen en la comunicación social? Para descubrirlo, me dediqué a observar a los niños en su ambiente escolar; Díaz eligió el laboratorio.

Con Ruth A. Garvin realicé el seguimiento de 36 niños apalaches de cinco a diez años, pertenecientes a familias de renta baja y que asistían a una escuela de la misión en las montañas del este de Kentucky. Registramos el habla en el aula, en el patio, en los pasillos y en el comedor durante todo el día. (Prestamos especial atención a los comentarios no dirigidos específicamente a un oyente.)

Rara vez se daba habla egocéntrica, el punto central de Piaget. Casi todos los comentarios que oímos o bien describían las acciones de un niño o bien servían para dirigirlas, lo que respaldaba la hipótesis según la cual la autoorientación constituye la función principal del soliloquio. Por otra parte, los niños hablaban consigo mismos más a menudo cuando trabajaban solos en tareas arduas y cuando sus profesores no podían ayudarles inmediatamente; esto es, cuando los pequeños tenían que tomar las riendas de su propio comportamiento.

Obtuvimos asimismo indicios de que el habla privada se desarrolla igual en todos los niños y que surge en la experiencia social. En el caso de los alumnos apalaches, el comportamiento fue cambiando paralelamente con el crecimiento, según un patrón muy similar al establecido por Kohlberg 15 años antes.

Los niños de clase media, como los observados por Kohlberg, hablan solos en voz alta con una frecuencia creciente entre los cuatro y los seis años de edad. Después, durante la escuela elemental, su soliloquio adopta la forma de una murmuración inaudible. Los niños apalaches siguieron, aunque con mayor lentitud, la misma pauta. A los 10 años, más del 40 por ciento de su habla privada continuaba siendo perceptible, frente al siete por ciento entre los niños de la misma edad estudiados por Kohlberg.

Para conocer las causas de esa discrepancia, investigamos la cultura apalache. Decubrimos un dato importante. Mientras que los padres de clase media conversan a menudo con sus hijos, los padres apalaches lo hacen con mucha menos frecuencia y, además, recurren a los gestos más que a las palabras. Si Vygotsky anda en lo cierto, el habla privada dimana de la comunicación social, de modo que ese taciturno ambiente familiar podría explicar el lento desarrollo del habla privada en los pequeños.

Variedades de habla privada

CATEGORIA	DESCRIPCION	EJEMPLO
Comunicación egocéntrica	Frases dirigidas a otro que carecen de sentido desde la perspectiva del oyente.	David le dice a Marcos, que está sentado a su lado en la estera, "se rompió", sin explicar qué o cuándo.
Juego de fantasía	Un niño representa cierto papel y habla a objetos o crea efectos sonoros para ellos.	Borja increpa "¡fuera de mi camino!" a una silla después de chocar con ella.
Liberación de las emociones	Comentarios no dirigidos a un oyente que expresan sentimientos o que parecen intentos de corregir sentimientos sobre acontecimientos o pensamientos pasados.	Raquel está sentada en su pupitre con una mirada ansiosa repitiendo para sí misma, "mi mamá está enferma, mi mamá está enferma".
Autodirección	Un niño describe la tarea que tiene en las manos y se da a sí mismo instrucciones en voz alta.	Gracia, mientras hace los ejercicios de su libro de matemáticas, dice en voz alta, "seis". Luego, contando con los dedos, continúa, "siete, ocho, nueve, 10. Es 10. La respuesta es 10."
Recitación en voz alta	Un niño lee en voz alta un escrito o pronuncia palabras.	"Sher-lock Holm-lock, Sherlock Holme,"Jan lee, quitando la "s" final en su segundo, y más satisfactorio, intento.
Murmuración inaudible	Murmuraciones que el observador no puede entender.	Olga murmura inaudiblemente para sí misma mientras intenta resolver un problema de matemáticas.

Mientras llevábamos a cabo este estudio, Díaz y Marnie H. Frauenglass grabaron en vídeo a 32 niños de tres a seis años que emparejaban imágenes y resolvían rompecabezas en el laboratorio. Observaron también que el habla privada se hace menos audible con la edad. Sin embargo, sus resultados, junto con los de otros investigadores, plantean varias objeciones a la teoría de Vygotsky. En primer lugar, muchos niños emitían sólo unas pocas expresiones, y algunos ninguna; contradecíase, pues, que el hablar solos tuviera carácter general.



2. LEV S. VYGOTSKY, con una de sus hijas. Sostuvo que el habla privada desempeña un papel positivo en el desarrollo infantil. Dadas las condiciones políticas reinantes en la Unión Soviética, sin embargo, su trabajo permaneció desconocido en Occidente hasta los años sesenta.

Surgió otra dificultad. Si el habla privada facilita, de acuerdo con Vygotsky, la autorregulación, debe estar relacionada con el comportamiento del niño mientras trabaja y con su rendimiento. Pero en el estudio de Frauenglass y Díaz los niños que recurrían más al soliloquio eran los que hacían peor las tareas. Otros trabajos hablaban también de asociaciones débiles, y a veces negativas, entre el habla privada y el rendimiento.

Díaz extrajo algunas ideas perspicaces de esos resultados. Tras un examen atento de la definición dada por Vygotsky de zona de desarrollo proximal, llegó a la conclusión de que las tareas inducidas en el laboratorio quizá no fueran las adecuadas para instar el habla privada en los niños. Algunos podrían estar tan familiarizados con la resolución de rompecabezas y emparejamiento de imágenes que las operaciones cognoscitivas necesarias para su ejecución serían ya automáticas. Para otros, sin embargo, habrían resultado tan difíciles, que no pudieron acometerlas sin ayuda. En ninguno de los dos casos cabría esperar que se produjera habla privada autodirectiva. Además, siguiendo el razonamiento de Díaz, dado que el recurso a hablar solos se apremia cuando los niños están en apuros, ese fenómeno coincidiría a menudo con el fracaso de la tarea. Sugirió que el efecto beneficioso del habla privada podría retrasarse.

Volviendo a las aulas (esta vez, a la escuela laboratorio de la Universidad estatal de Illinois), me embarqué en una serie de estudios para comprobar tales hipótesis. Mis colaboradores anotaron puntillosamente los soliloquios y las acciones relacionadas con las tareas encomendadas de 75 alumnos de los grados primero a tercero, relativas a la resolución, ellos solos en sus pupitres, de problemas de matemáticas. Los maestros consideraban que el trabajo era el adecuado para todos los niños. Con Jennifer A. Bivens me encargué del seguimiento de los alumnos de primero, y controlamos el comportamiento de los mismos en segundo y tercero.

Todos hablaron consigo mismos un promedio del 60 por ciento del tiempo. Según se había observado en estudios precedentes, muchos niños cuyas manifestaciones describían o comentaban su actividad, obtuvieron calificaciones menores en los deberes y en los exámenes de aprovechamiento ese mismo año. Sin embargo, el habla privada que era típica de una edad concreta anunciaba logros en matemáticas con el paso del tiempo. En particular, los alumnos de primer grado que hicieron muchos comentarios autodirectores, en voz alta o baja, progresaron en las matemáticas de segundo grado. De igual modo, los de segundo que murmuraron a menudo comprendían con más facilidad las de tercer grado el año siguiente.

Además, la relación que observamos entre el recurso al soliloquio y su rendimiento en la tarea respaldaba la hipótesis de Vygotsky según la cual los comentarios autoorientativos ayudan a dirigir sus acciones. Los niños cuya habla incluía abundancia de palabras carentes de importancia para la tarea o mucha expresión emocional solían revolverse en sus asientos, morder el extremo del lapicero o golpearlo contra el pupitre. Los niños que hacían frecuentes comentarios audibles sobre su trabajo utilizaban, en ayuda para superar sus dificultades, más técnicas no verbales: contar con los dedos o seguir una línea de texto con un bolígrafo. Los niños, por último, que más a menudo usaban el habla privada en voz baja rara vez se impacientaban y eran muy atentos.

En general, los niños que progresaban más deprisa desde los comentarios audibles hacia el habla interna eran también los más adelantados en capacidad para controlar la actividad motriz y centrar la atención. El desarrollo del habla privada y el comportamiento relacionado con la tarea desarrollada iban, pues, de la mano.

En una investigación posterior, Sarah T. Spuhl y la autora se propusieron comprobar en el laboratorio la relación dinámica que Vygotsky atisbó entre el habla privada y el aprendizaje; a saber, los soliloquios disminuyen conforme aumenta el rendimiento. Añadimos además una nueva dimensión a nuestro trabajo: averiguar en qué medida la interacción entre el niño y el adulto puede fomentar la autorregulación a través del habla privada.

Pedimos a 30 niños de cuatro a cinco años que montaran cierto modelo con piezas de lego. Cada niño intentó el ejercicio en sesiones de 15 minutos, programadas con una diferencia entre ellas de no más de dos a cuatro días. Esta temporización nos permitió seguir la trayectoria de su habilidad creciente. Incorporamos una innovación: realizamos una evaluación previa de cada pequeño para garantizar que las tareas que se les iba a proponer fueran estimulantes. Sólo participaron constructores noveles en el juego. Dos semanas antes de que empezaran las sesiones, grabamos en vídeo la ayuda que cada madre prestaba a su hijo en actividades que precisaban aptitudes similares a las necesarias en dichos montajes; por ejemplo, encajar bloques y emparejar colores y formas.

Examinamos luego la comunicación establecida entre madre e hijo conforme iban resolviendo juntos los problemas. De acuerdo con la investigación anterior, los padres cariñosos e interesados, que orientan y alientan al niño a adquirir nuevas habilidades sin exagerar el control, fomentan la capacitación. (Los psicólogos denominan paternidad y maternidad autoritativas a esa relación.) Por contra,

los padres autoritarios (poco afecto y mucho control) y los permisivos (mucho afecto y poco control) son un anuncio de problemas en el aprendizaje y la idoneidad. Basándonos en tales pruebas, pensamos que el estilo autoritativo captaría mejor las características de la enseñanza impartida por el adulto que deseábamos identificar.

De acuerdo con los datos obtenidos, los niños con madres autoritativas usaban con más frecuencia el soliloquio de autoorientación. Entre los niños de cuatro años, el progreso en las tres sesiones de construcción del lego resultó ser mayor en los que experimentaban una enseñanza autoritativa. Además, hicimos un análisis estadístico especial, del que se infería que el habla privada promedia la relación entre paternidad-maternidad autoritativa y éxito de la tarea: descubrimiento compatible con lo supuesto por Vygotsky.

A diferencia de la anterior investigación de laboratorio, todos los niños de nuestra muestra hablaban solos. Según lo esperado, los comentarios infantiles iban interiorizándose en el transcurso de las tres sesiones a medida que aumentaban las habilidades de los niños para armar el lego. Y, una vez más, el soliloquio permitía predecir los futuros progresos mejor que el éxito obtenido en la tarea concurrente. En particular, los niños que usaban el habla privada adecuada para su edad (expresiones audibles y autodirectivas a los cuatro años y murmullos audibles a los cinco) avanzaban con mayor seguridad.

A continuación me interesé por los niños con problemas graves de aprendizaje y de conducta. Muchos psicólogos sostenían que los alumnos de primaria distraídos, impulsivos y retrasados distinguíanse también por el escaso uso del soliloquio. Para tratar a esos niños, se habían diseñado y puesto en práctica programas de formación orientados a fomentar el habla privada; se les pide, por ejemplo, que imiten al terapeuta, quien simula el habla privada autodirectiva mientras realiza una tarea. Luego, el psicólogo sólo mueve los labios. Por fin, los niños son invitados a imitarle disimuladamente.

Pero semejante planteamiento, no obstante su ingenioso diseño, fracasó casi siempre. Sospeché que se había actuado con precipitación. Antes, debíase haber investigado la forma en que los niños con problemas de aprendizaje y de conducta recurren al habla privada. Seguían sin analizarse las expresiones espontáneas autorreguladoras de esos niños.

Para cubrir ese vacío, estudié con Michael K. Potts 19 niños de seis a doce años que padecían trastornos de hiperactividad con déficit de atención (THDA), afección caracterizada por ausencia de atención, comportamiento impulsivo e hiperactividad. Registramos el habla privada mientras se aprestaban a resolver en el pupitre problemas de matemáticas. Comparamos esas observaciones con los soliloquios de 19 niños normales de pareja edad y capacidad verbal similar.

En negación de lo supuesto al aplicar el autoentrenamiento, los niños con THDA no eran deficitarios en el recurso al soliloquio. Antes bien, hacían más afirmaciones autodirectivas y audibles que los niños normales. Además, vimos, al abordar las tendencias relacionadas con la edad, que la única diferencia entre los dos grupos residía en que los niños con THDA realizaban la transición de habla audible a formas interiorizadas a una edad posterior.

Dimos con una explicación de esa demora: el grave déficit de atención de los niños con THDA impedía que su habla privada lograra un control eficaz sobre su comportamiento. En primer lugar, sólo se encontró una relación entre habla audible y mayor atención a las tareas de matemáticas en los niños con THDA que se distraían menos. En segundo lugar, hicimos un seguimiento de una submuestra de sujetos con THDA en dos situaciones: mientras recibían estimulantes, la terapia farmacológica habitual en ese trastorno, y cuando no estaban sujetos a ese tratamiento. (Aunque los estimulantes no curan el THDA, sí potencian la atención y el rendimiento académico en la mayoría de los niños que los toman.) Observamos que estos fármacos provocaban la maduración drástica del habla privada en los niños con THDA. Y sólo cuando esos niños eran sometidos a tratamiento farmacológico, su forma más madura, el murmullo inaudible, se relacionaba con un mejor autocontrol.

El cariz prometedor de estos descubrimientos me alentó a incluir niños con dificultades de aprendizaje en la investigación. Con Steven Landau estudié 112 alumnos de tercero a sexto grado que realizaban ejercicios de matemáticas y de inglés. La mitad de los niños entraban de plano en el grupo de discapacitados para el aprendizaje, es decir, el formado por chicos cuyos resultados académicos están muy por debajo de sus dotes intelectuales. La otra mitad sirvió de control. Lo mismo que en el estudio del THDA, comprobamos que los discapacitados para el aprendizaje utilizaban más expresiones de autodirección y audibles e interiorizaban su habla privada a una edad posterior que los niños sin esas dificultades. Cuando observamos un subgrupo de los primeros con síntomas, además, de THDA, esta tendencia se intensificaba.

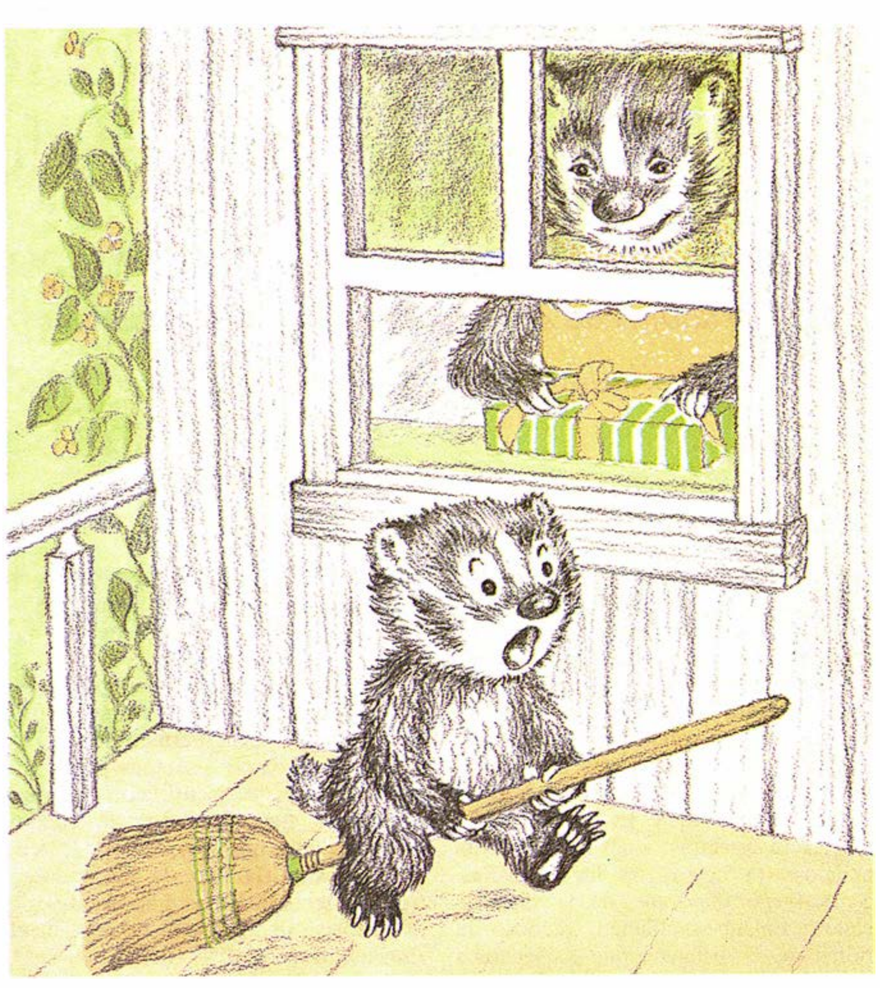
La investigación realizada con niños que adolecían de dificultades de



3. UNA ORIENTACION HABIL da a los niños el andamiaje que precisan para usar con eficacia el habla privada. Ante un niño enfrentado a una tarea desafiante, el adulto puede ayudarle con instrucciones o estrategias habladas que contribuyen a su éxito. El niño puede incorporar luego el lenguaje de esas conversaciones en su soliloquio y utilizarlo más tarde para guiar su propio esfuerzo.

aprendizaje corrobora la opinión de Vygotsky sobre el habla privada: siguen el mismo curso de desarrollo que sus compañeros de la misma edad no afectados; ahora bien, por culpa de un procesamiento cognoscitivo deterioriado y una mermada capacidad para prestar atención, les resultan más difíciles las tareas académicas, dificultad que, a su vez, complica la autorregulación verbal. Por lo que hemos podido comprobar, el ejercitar a hablar solos a niños con problemas de aprendizaje y de comportamiento, mientras realizan tareas cognoscitivas, no es más que demandarles una habilidad que ya poseen; las intervenciones encaminadas a acelerar el proceso de autocomunicación silenciosa podrían ser incluso contraproducentes. Cuando se concentran, los alumnos con THDA y los tardos en el aprendizaje se agarran al soliloquio audible en busca de apoyo compensatorio de su discapacidad cognoscitiva.

¿Cómo aprovechar los avances obtenidos en el conocimiento del habla privada en beneficio de una mejor pedagogía con niños normales y con niños que sufren problemas de aprendizaje y conducta? Tomados en su conjunto los resultados de las investigaciones, se advierte que el soliloquio constituye una herramienta universal de resolución de problemas para los niños que crecen en ambientes enriquecedores y socialmente interactivos. Algunos factores interdependientes (dureza de una tarea, contexto social y características individuales del niño) determinan el grado y soltura con que cada uno echa mano de los soliloquios para guiar su comportamiento. No se trata, pues, de considerar el habla privada una habilidad en la que haya que ejercitarse, sino de crear las condiciones que ayuden a emplearla con eficacia.



4. NUESTRO PEQUEÑO TEJON es protagonista, en muchos lugares, de aventuras infantiles. En la escena lo vemos jaleándose a sí mismo para disimular los celos que siente hacia su hermana pequeña, que recibe el regalo de su primer cumpleaños. El soliloquio es un recurso común en la literatura infantil.

Cuando un niño emprende una nueva tarea, necesita el apoyo cordial de un adulto paciente y alentador, que le ofrezca la ayuda proporcionada a las capacidades del pequeño. Si no acaba de entender en qué consiste la tarea, por ejemplo, el adulto podría empezar por suministrarle orientaciones explícitas. Una vez comprendida la relación entre esas acciones y el fin de la tarea, podría sugerirle estrategias para conseguir el objetivo. Conforme el niño se vaya adueñando de sus propias iniciativas, el adulto se irá retirando.

Con sobrada frecuencia, a los niños distraídos e impulsivos se les niega la construcción de este andamiaje para el aprendizaje. En la atmósfera de tensión que se crea entre esos niños y los adultos, aquéllos sufren órdenes, reprimendas y críticas, justo lo que les impide aprender cómo controlar sus propias acciones.

Por último, padres y profesores deben ser conscientes del valor funcional del hablar solos. Hemos explicado que se trata de un comportamiento sano, adaptativo y decisivo, que algunos niños necesitan utilizarlo más a menudo y durante períodos más largos que otros. Pero son todavía muchos los adultos que siguen considerando el habla privada como una conducta sin sentido, socialmente inaceptable, delatora incluso de alguna debilidad mental; e intentan, en consecuencia, disuadirles de tal práctica.

La verdad es que, en casa, los padres pueden escuchar los soliloquios de sus hijos y adquirir así mayor información sobre sus planes, metas y dificultades. De igual modo, los profesores han de ser conscientes de que, cuando los alumnos recurren en demasía para su edad al habla privada, están demandando apoyo y orientación suplementarios.

Queda mucho por descubrir sobre el uso que los niños hacen del habla privada espontánea para resolver los problemas. Gracias a la teoría de Vygotsky hemos progresado muchísimo en la comprensión de este fenómeno. En ella nos basamos para diseñar una pedagogía más eficaz.

BIBLIOGRAFIA COMPLEMENTARIA

DEVELOPMENT AND FUNCTIONAL SIGNIFICANCE OF PRIVATE SPEECH AMONG ATTENTION-DEFICIT HIPERACTIVITY DISOREDED AND NORMAL BOYS. Laura E. Berk y Michael K. Potts en *Journal of Abnormal Child Psychology*, vol 19, n.º 3, págs. 357-377; junio de 1991.

PRIVATE SPEECH: FROM SOCIAL INTERACTION TO SELF-REGULATION. Dirigido por Rafael M. Díaz y Laura E. Berk. Lawrence Erlbaum Associates, 1992.

PRIVATE SPEECH OF LEARNING DISABLED AND NORMALLY ACHIEVING CHILDREN IN CLASSROOM ACADEMIC AND LABORATORY CONTEXTS. Laura E. Berk y Steven Landau en *Child Development*, vol. 64, n.º 2, páginas 556-571; abril de 1993.

Una resolución de las paradojas de Zenón

A lo largo de milenios, matemáticos y filósofos se han esforzado por refutar las paradojas de Zenón, un conjunto de aserciones sorprendentes que niegan la posibilidad del movimiento. Por fin se ha obtenido una solución

William I. McLaughlin

Sucedió una vez que Aquiles, héroe de velocidad legendaria, se encontró con una tortuga en su camino. La tortuga, que tenía la mente mucho más rápida que los pies, le retó a una carrera. Divertido, Aquiles aceptó. La tortuga solicitó cierta ventaja, pues a la vista estaba que era mucho más lenta que el semidiós. Aquiles aceptó encantado, y la tortuga inició la carrera mientras aquél se entretenía atándose la correa de las sandalias. Salió luego disparado de la raya de salida. En un santiamén, recorrió la mitad de la distancia que lo separaba de la tortuga. En otro abrir y cerrar de ojos había recorrido tres cuartas partes del intervalo. En otro instante, había cubierto ya los siete octavos y, en otro, los quince dieciseisavos. Pero, por muy aprisa que corriese, siempre había entre el pélida y la tortuga una distancia que salvar. Daba la impresión de que jamás alcanzaría al quelonio perseverante.

Si Aquiles hubiese dedicado menos tiempo a la gimnasia y más al estudio de la filosofía, habría sabido que estaba protagonizando el clásico ejemplo utilizado por Zenón para ilustrar una de sus paradojas, la que sostiene la imposibilidad de cualquier movimiento. Zenón ideó la paradoja de Aquiles y la tortuga y otros enigmas similares para respaldar las teorías filosóficas de Parménides, su maestro.

Procedían de la colonia griega de Elea, situada en la Italia meridional. Hacia el 445 a.C., Parménides y Zenón se trasladaron a Atenas para visitar a Sócrates y discutir con él sobre cuestiones fundamentales de filosofía. El acontecimiento, que, si se realizó, constituiría uno de los encuentros intelectuales más importantes de la historia, se halla consignado en el *Parménides*.

En este diálogo de Platón, aparece Parménides, un pensador ilustre de 65 años, presentándole al joven Sócrates una tesis sobrecogedora: la "realidad" es una entidad singular sin cambios ni fisura alguna en su unidad. El mundo físico, argumentaba, es monolítico. En particular, resulta imposible el movimiento. Si bien el rechazo de la pluralidad y el cambio parece asunto subjetivísimo, se trata de un planteamiento que, expuesto en términos generales, ha merecido desde entonces la atención de muchos estudiosos. Así, por ejemplo, el "idealismo absoluto" del filósofo de Oxford F. H. Bradley (1846-1924) tiene puntos de contacto con la perspectiva de Parménides.

Semejante descripción del mundo contradice nuestra experiencia cotidiana y relega nuestras percepciones básicas al reino de la ilusión. Para defender su tesis, Parménides se apoyaba en poderosos argumentos de Zenón, razonamientos que Aristóteles incorporaría en sus obras. A lo largo de casi dos milenios y medio, las paradojas de Zenón han sido motivo de controversia y objeto de análisis. Hoy, por fin, podemos resolver las paradojas gracias a cierta formulación del cálculo matemático desarrollada en los últimos diez años, más o menos. La resolución en cuestión pivota sobre la noción de infintésimo, conocida desde hace tiempo pero mirada siempre con escepticismo por distintos pensadores.

El relato de Aquiles y la tortuga describe "la dicotomía", una de las paradojas de Zenón: cualquier distancia que deba ser recorrida por un objeto, la que separa, por ejemplo, ambos contendientes, puede ir dividiéndose por la mitad (1/2,1/4,1/8, etcétera) en una infinidad de segmentos, cada uno de los cuales representa un espacio por recorrer. De ello se infiere, dice Zenón, que no puede llevarse a cabo ningún movimiento, pues siempre queda una cierta distancia por recorrer, sin que importe cuán pequeña sea.

Conviene advertir que Zenón no afirma que sea imposible sumar una infinidad de longitudes y obtener una distancia finita (un simple vistazo a la geometría de una línea dividida en infinidad de trozos muestra de inmediato, sin cálculos refinados, que una infinidad de piezas suman un intervalo finito). Antes bien, la objeción de Zenón contra la idea de movimiento basa su fuerza en la obligación de explicar cómo completar, en una serie, un número infinito de actos (los que hay que realizar al cruzar un intervalo).

Zenón lanzó una segunda andanada contra los fundamentos conceptuales del movimiento. Repetía la primera argumentación desde un enfoque algo distinto. Así reza, en breve, la segunda paradoja: antes de que una flecha (o cualquier otro objeto) alcance el punto medio de su hipotético recorrido (hecho que en el caso precedente se daba por supuesto) es preciso que cubra una cuarta parte de dicha distancia, y antes una octa-

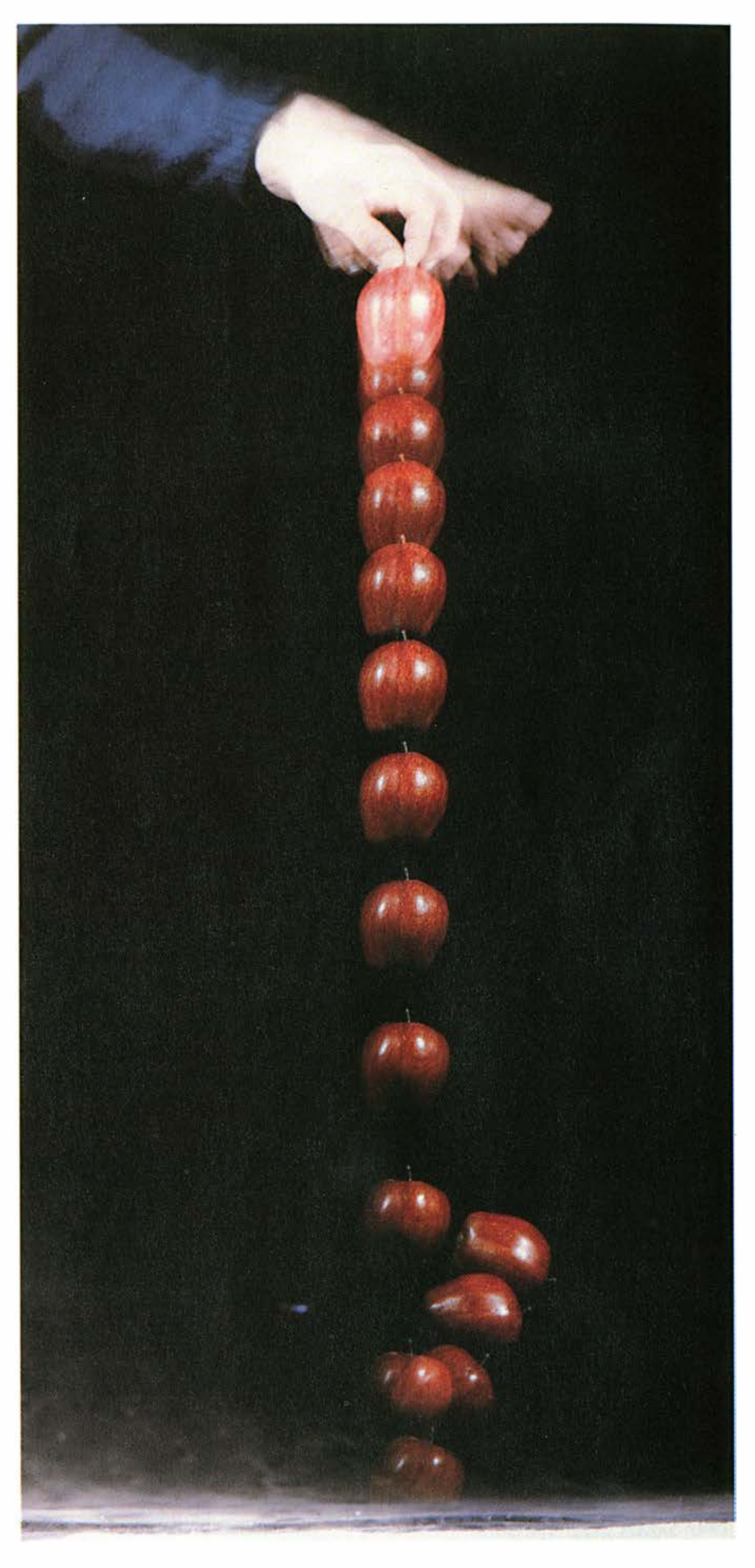
WILLIAM I. McLAUGHLIN es experto en astrofísica avanzada del Laboratorio de Propulsión a Chorro de Pasadena, donde trabaja desde 1971. Ha participado en muchos proyectos del programa espacial de los Estados Unidos; entre otros, el alunizaje del Apolo, la misión Viking a Marte, el Satélite Astronómico de Infrarrojos (IRAS) y el proyecto Voyager. Tras licenciarse en ingeniería eléctrica en 1963, se doctoró en matemáticas por la Universidad de California en Berkeley. McLaughlin investiga, además, en epistemología.

va, etcétera. Es decir, lo mismo que en la primera paradoja, el razonamiento puede continuarse indefinidamente llevándonos a una regresión infinita, y, por consiguiente, a su tesis: el movimiento no pudo nunca iniciarse.

En la tercera paradoja, Zenón toma un rumbo distinto. Afirma que el propio concepto de movimiento carece de contenido. Zenón nos invita a considerar la flecha en un instante arbitrario de su trayectoria. En dicho momento, la flecha ocupa una región del espacio que es igual a su propia longitud, y no evidencia movimiento alguno. Y puesto que esta observación es válida en cada instante, la flecha no se mueve. Desde un punto de vista histórico se trata de la objeción que mayores quebraderos de cabeza ha creado en todos los comentaristas que se han acercado a las paradojas de Zenón.

Muchos son, en efecto, los filósofos y matemáticos que se han esforzado por desanudar las paradojas del eleata. Los enfoques más directos se han limitado a negar simplemente la existencia de problema alguno. Así, en 1782 Johann Gottlieb Waldin, un profesor de filosofía alemán, escribió que Zenón, al argumentar en contra del movimiento, daba ya por supuesta su existencia. Evidentemente el filósofo germano no parecía familiarizado con el recurso argumentativo de reducción al absurdo, en la que, partiendo de la aceptación de cierto estado de cosas, se establece luego que dicha admisión desemboca en una conclusión ilógica.

Otros estudiosos, sin embargo, consiguieron ciertos progresos cuando abordaron la cuestión sobre la posibilidad de que, en el mundo físico, se dé un número infinito de acciones. En sus explicaciones se ha cruzado continuamente la idea de infinitésimo, intervalo de espacio o



1. ¿CAE LA MANZANA? Razonaría Zenón que, puesto que en cada instante la manzana se nos muestra congelada en el espacio, nunca se halla en movimiento. Además, afirmaría que no hay forma alguna de probar que la manzana alcanzará el suelo. Antes de alcanzarlo, debería descender la mitad de la distancia que separaba la mano del hombre del suelo; y antes, debería recorrer la mitad de la distancia que falta, y nuevamente la mitad de ésta, y así sucesivamente. Siempre habrá cierto espacio entre la manzana y el suelo. Mediante una argumentación lógica parecida, Zenón se cuestiona incluso la posibilidad de que la manzana inicie siquiera el movimiento.



2. LA CARRERA entre Aquiles y la tortuga ilustra una de las paradojas de Zenón. Aquiles concede cierta ventaja al quelonio. El pélida se ve obligado a recorrer la mitad de la distancia que los separa, después las tres cuartas partes, luego las siete octavas partes y así sucesivamente hasta el infinito. Parece, pues, evidente, que nunca podrá alinearse con el parsimonioso animal.

tiempo que encarna la quintaesencia de la pequeñez. Un cantidad infinitesimal, una vez aceptada, debería ser tan próxima a cero que resultaría impotente numéricamente hablando. Tales cantidades escaparían a toda medición, por precisa que fuese, cual arena a través de una criba.

Giovanni Benedetti (1530-1590), precursor de Galileo, postuló que, cuando en Zenón un objeto aparece congelado en el espacio, el filósofo sólo observa una parte de la acción, como si viéramos un fotograma en vez de la película entera. Entre las imágenes estáticas de Zenón, añadía, hay instantes infinitamente pequeños en los que el objeto recorre distancias igualmente pequeñas.

Otros eludían la cuestión aduciendo que, en el mundo físico, los intervalos no pueden subdividirse infinidad de veces. Friedrich Adolf Trendelenburg (1802-1872), de la Universidad de Berlín, elaboró todo un sistema filosófico que explicaba la percepción humana a través del movimiento. De ese modo se libraba de tener que dar razón de éste.

Un camino parecido siguió, en nuestro siglo, el filósofo y matemático inglés Alfred North Whitehead (1861-1947). Construyó un sistema metafísico basado en el cambio, con el movimiento como un caso especial. Whitehead fundaba su respuesta a las paradojas de Zenón en la extensión que, necesariamente, han de tener los sucesos del mundo físico; vale decir: tales fenómenos no pueden ser puntuales. Por su parte, el filósofo escocés David Hume (1711-1776) escribió que "todas las ideas de cantidad sobre las que razonan los matemáticos son particulares y, en cuanto tales, sugeridas por los sentidos y la imaginación, y, por consiguiente, no pueden ser infinitamente divisibles".

En otras palabras, la cuestión de los infinitesimales (y de si existen o no) generó por sí sola una amplia y bronca bibliografía. Hasta ayer mismo, los matemáticos estaban convencidos de que se trataba de una simple quimera. El obispo irlandés George Berkeley (1685-1753), famoso por su teoría idealista que niega la realidad de la materia, se enfrentó a los infinitesimales. Creía que los matemáticos de su época, incluido el propio Newton, no acertaban a entenderlos. "No son cantidades finitas, ni tampoco cantidades infinitamente pequeñas, pero tampoco son nada. ¿No podríamos llamarlos fantasmas de magnitudes que han abandonado este mundo para siempre?" Y señalaba: "Si los matemáticos pueden pensar en las fluxiones [las razones del cambio], en el cálculo diferencial y en cosas semejantes, entonces un poco de reflexión les mostrará que, al trabajar con tales métodos, no imaginan ni piensan en líneas o superficies menores que las que son perceptibles por los sentidos."

A los matemáticos les costaba acomodar los infinitésimos en sus descubrimientos, dejando aparte el recelo que les provocara en teoría. Creen algunos historiadores que el gran Arquímedes (aprox. 287-212 a.C.) logró obtener algunos de sus resultados matemáticos utilizando infinitesimales, pero en sus exposiciones públicas empleaba formas más tradicionales. Los infinitesimales dejaron su huella en los siglos XVII y XVIII, en el desarrollo del cálculo diferencial e integral. Los libros de texto han apelado reiteradamente a los "infinitesimales prácticos" para transmitir a los alumnos ciertas ideas del cálculo.

Cuando los analistas se propusieron hallar una justificación rigurosa de la existencia de estas cantidades exiguas, toparon con enormes dificultades. Andando el tiempo, los matemáticos del siglo XIX inventaron un sustituto técnico de los infinitésimos: la teoría de límites. Tal fue su éxito, que algunos matemáticos hablaron del "fin" de los infinitesimales en su disciplina. Pero, llegados los sesenta de nuestro siglo, la huella fantasmal de los infinitesimales se hizo cada vez más patente en los pasillos matemáticos, gracias a la obra de Abraham Robinson, lógico de la Universidad de Yale. Desde entonces, además del método de Robinson, se han ideado otros que echan mano de esas entidades.

Cuando Sylvia Millar y el autor de este artículo empezaron a trabajar sobre las paradojas de Zenón, contaban con la ventaja de que los infinitesimales se habían ganado un puesto de respeto en el universo matemático. Nos atraía su estudio porque, a nuestro parecer, aportaban un punto de vista microscópico de los detalles del movimiento. Edward Nelson, de la Universidad de Princeton, había creado la herramienta que se nos antojaba adecuada para nuestra investigación: la teoría de conjuntos internos (TCI), una rama especial de análisis no-estándar.

El método de Nelson permite interpretaciones sorprendentes de estructuras matemáticas que podrían reputarse familiares. Los resultados recuerdan, en su rareza, a las estruc-

turas de la teoría cuántica y de la relatividad general. Y si tenemos presente que estas dos teorías han requerido de un siglo casi entero para lograr una amplia aceptación, habrá que admirar el poder de la imaginación de Nelson.

Nelson adoptó un método novedoso para definir los infinitesimales. Los matemáticos acostumbran generalizar los sistemas numéricos ya existentes añadiéndoles objetos que tienen propiedades deseables para ellos, de forma muy parecida a la utilizada para sembrar fracciones entre enteros. Normalmente, el sistema numérico empleado en la matemática moderna crece, cual arrecife de coral, por acreción sobre un soporte: "Dios hizo los enteros, el resto es obra del hombre", sentenciaba Leopold Kronecker (1823-1891). Pero el camino seguido en la TCI consiste en "mirar" más a fondo los sistemas numéricos ya existentes y comprobar que alojan ya en su seno números que, de forma muy razonable, pueden ser considerados infinitesimales.

En rigor técnico, Nelson halla números no-estándar en la recta real añadiendo tres reglas, o axiomas, al conjunto de la decena de postulados que dan soporte a la mayoría de los sistemas matemáticos. (La teoría de conjuntos de Zermelo-Fraenkel es una de tales fundamentaciones matemáticas.) Estos añadidos introducen un término nuevo, el término "estándar", y nos ayudan a precisar cuál de nuestros antiguos conocidos del sistema numérico es estándar y cuál no-estándar. Los infinitesimales caen en la categoría no-estándar, junto con algunos otros números de los que hablaremos más adelante.

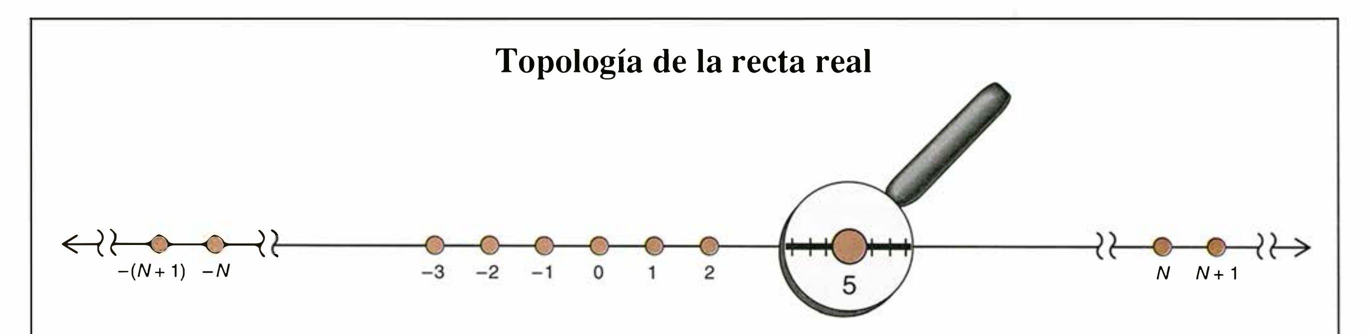
Para Nelson, infinitesimal es el número que se halla entre el cero y cualquier otro número estándar positivo. Podría pensarse, de entrada, que esto no nos lleva a la idea de pequeñez, pero los números estándar incluyen cada uno de los números concretos (y algunos otros) que podemos escribir en un papel o generar en una computadora: 10, pi, $^1/_{1000}$, etcétera. Por consiguiente, un infinitesimal es mayor que cero y menor que cualquier número estándar imaginable, por pequeño que sea. No resulta en absoluto evidente que tales infinitesimales existan, pero la validez conceptual de la teoría de conjuntos internos ha sido establecida con rigor parejo al que justifica nuestra creencia en otros sistemas matemáticos.

Los infinitesimales son entidades verdaderamente esquivas. Carácter que obedece a un dato matemático: aquel según el cual dos números concretos —de los que tienen contenido numérico— no pueden diferir jamás en una cantidad infinitesimal. La demostración, por reducción al absurdo, es fácil: la diferencia aritmética de dos números concretos debe ser concreta (y, por consiguiente, estándar). Si fuese infinitésima esa diferencia, la definición de infinitesimal como algo que es menor que cualquier número estándar se violaría. De ello se infiere que los extremos de un intervalo infinitesimal no podrán rubricarse jamás mediante números concretos. Por consiguiente, jamás podrá aprehenderse un intervalo infinitesimal a través de mediciones; los intervalos infinitesimales trascienden el dominio de la observación.

¿De qué manera podemos, pues, recurrir a esos números fantasma para resolver las paradojas de Zenón? De la exposición anterior se desprende que los puntos del espacio y del tiempo marcados con números concretos son puntos aislados. Una trayectoria y el intervalo temporal asociado a la misma se hallan densamente empaquetados con regiones infinitesimales. Gracias a ello podemos aceptar la tercera objeción de Zenón: la punta de la flecha es captada "estroboscópicamente" en reposo en ciertos puntos del tiempo que están indicados de forma precisa, pero en la mayor parte de la extensión acontece alguna suerte de movimiento. Sin embargo, tal movimiento es inmune a la crítica zenoniana porque, de acuerdo con los postulados impuestos, se ve forzado a ocurrir en los segmentos infinitesimales. Y su inefabilidad le proporciona cierta mampara de protección o filtro.

¿Cabría que el proceso del movimiento en un intervalo de ésos consistiera en un avance uniforme a lo largo del intervalo o en un salto instantáneo de un extremo del intervalo infinitesimal al siguiente? ¿Debe, acaso, el movimiento comprender una serie de etapas intermedias o quizás un proceso que se halla, a la vez, fuera del tiempo y del espacio? Las posibilidades son infinitas, y ninguna de ellas podrá verificarse ni desecharse porque nunca podrán con-

Topología de la recta real



Los números reales constan de los enteros (positivos y negativos), los números racionales (que se expresan por medio de una fracción) y los números irracionales (que no pueden expresarse por medio de una fracción). Los números reales se representan por medio de los puntos de la recta real (*figura superior*).

El matemático Edward Nelson, de la Universidad de Princeton, etiquetó, en este sistema estándar de números, tres tipos de números no-estándar. Por números no-estándar infinitesimales se entienden los que son menores que cualquier número estándar positivo y a la vez mayores que cero. Los números no-estándar mixtos, que se presentan hacinados en torno de un entero (por ejemplo, el cinco), resultan de sumar y restar cantidades infinitésimas a los números estándar (el cinco, por ejemplo). De hecho cada número estándar se halla rodeado por dichos números mixtos, formando su entorno no-estándar. Los números no-estándar ilimitados, representados por N y $N+1$, son los inversos de los números infinitesimales no-estándar. Cada uno de los números no-estándar ilimitados es mayor que cualquier número estándar y, a su vez, menor que los números reales infinitos. Los números reales no-estándar se muestran muy útiles para resolver las paradojas de Zenón.

trolarse los intervalos infinitesimales. Se impone reconocer el mérito del rechazo a Benedetti, Trendelenburg y Whitehead por su perspicacia, perspicacia que hoy podemos formalizar con la TCI.

Resulta más fácil dar respuesta a las dos primeras paradojas que a la tercera. Mas, para hacerlo, hemos de apelar a otro recurso matemático de la teoría de conjuntos internos. Cualquier conjunto infinito de números contiene un número no-estándar. Antes de diseñar las implicaciones zenonianas de esta sentencia, es necesario hablar de los otros tipos de números no-estándar que se generan a partir de los números infinitesimales.

Empecemos por considerar la totalidad de los números infinitesimales, que, por definición, se insieren entre el cero y cualquier número estándar positivo. Coloquemos, luego, delante de cada uno de ellos, un signo menos. Disponemos así de un racimo formado por estos minúsculos objetos, simétrico con respecto del cero al de los números infinitesimales. Para crear números no-estándar "mixtos", tomamos un número estándar, por ejemplo, el número un medio, y le añadimos cada uno de los infinitesimales no-estándar que se hacinan en torno al cero. Este proceso aditivo traslada el racimo original de los infinitesimales a posiciones que se hallan situadas a un lado y otro del número un medio. De manera similar, cada número estándar puede considerarse rodeado de sus vecinos, todos ellos números no-estándar, que se encuentran a una distancia infinitesimal de aquél.

El tercer tipo de números no-estándar está formado por los inversos de los infinitesimales. Puesto que un infinitesimal es muy pequeño, su inverso será muy grande (en el reino de lo estándar, el inverso de una millonésima es un millón). Este tipo de números no-estándar recibe el nombre de número ilimitado. Los números ilimitados, si bien muy grandes, son finitos y, por tanto, menores que los números reales verdaderamente infinitos creados por los matemáticos. Los números ilimitados viven en una zona crepuscular situada entre los números estándar familiares, que son finitos, y los infinitos.

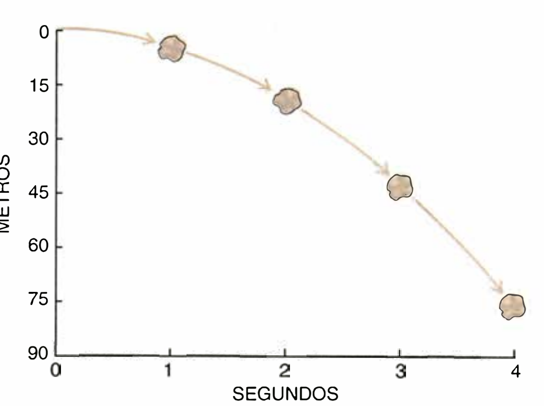
Entonces si, tal como se establece en la TCI, cada conjunto infinito contiene números no-estándar, resulta que la serie infinita de puntos-observados, utilizados por Zenón para medir el movimiento en su primera paradoja, debe contener un número no-estándar que es mixto. De hecho, dado que la serie infinita de Zenón repta en aproximación creciente hacia la unidad, uno de los miembros de la serie terminará por hallarse a una distancia infinitesimal de uno. A partir de este punto, todos los miembros de la serie que le siguen serán miembros no-estándar del racimo del uno, y ni nuestro presocrático de Elea ni ningún filósofo contemporáneo de hoy podría observar el avance de un objeto móvil cuando se encuentra en esta región inaccesible.

Resulta un tanto curioso que se acuda al infinito —la supuesta arma de Zenón— para vaciar de contenido sus propias afirmaciones. Para resolver su primera paradoja, sólo precisamos recurrir al principio epistemológico que establece que no somos responsables de no poder dar explicaciones de aquellas situaciones que no podemos observar. La serie infinita de puntos-observados de Zenón contiene números no-estándar, los cuales carecen de significación numérica; por consiguiente, podemos rechazar su razonamiento basándonos precisamente en tales entidades.

Puesto que nadie puede, ni siquiera en principio, contemplar el dominio completo de los puntos-observa-

Cálculo por medio de infinitesimales

Para comprender la relación que existe entre los infinitesimales y el cálculo diferencial, consideremos la caída de una piedra. La distancia que recorre la piedra medida en metros se puede calcular por medio de la fórmula $s = 4.9t$, en donde t indica el tiempo transcurrido medido en segundos: si la piedra ha caído durante dos segundos, habrá recorrido 9,8 metros.



Supongamos, no obstante, que se quiere calcular la velocidad instantánea de la piedra. La velocidad media de un objeto que se halla en movimiento es igual a la distancia recorrida dividida por la cantidad total de tiempo que ha precisado para hacerlo. Aplicando esta fórmula a un cambio infinitesimal en la distancia total y en el tiempo, podemos obtener una muy buena aproximación de la velocidad instántanea del objeto.

Supongamos que dt representa un cambio infinitesimal en el tiempo y ds un cambio infinitesimal en la distancia. El cómputo para obtener la velocidad instantánea de la piedra después de transcurrido un segundo en su movimiento, será entonces la siguiente: el intervalo de tiempo en consideración va de $t = 1$ a $t = 1 + dt$. La posición de la piedra durante este tiempo cambia de $s = 4,9(1)^2$ a $s = 4,9(1 + dt)^2$. El cambio total de la distancia, $9,8dt + 4,9dt^2$, dividido por dt, es la velocidad media que buscábamos, $9,8 + 4,9dt$.

Ahora bien, puesto que $4,9dt$ es una cantidad infinitesimal que resulta indetectable a cualquier intento e irrelevante para nuestros propósitos, se puede considerar igual a 0. Así pues, tras un segundo de recorrido, la fórmula nos dice que la velocidad instantánea de la piedra que cae es de 9,8 metros por segundo.

Esta manipulación, no obstante, se asemeja a la que se utiliza en el cálculo diferencial clásico. En él, el residuo $4,9dt$ no puede ser eliminado al final del cálculo; se trata de una cantidad que no es infinitesimal. Se impone, en dicho cálculo, acudir a la teoría de límites. En esencia, el proceso del límite hace que el intervalo dt sea suficientemente pequeño para que la velocidad media se acerque, arbitrariamente, a 9,8. Lo mismo que antes, la velocidad instantánea de la piedra tras un segundo de recorrido es igual a 9,8 metros por segundo. De forma análoga, un uso juicioso de las regiones infinitesimales facilita el cálculo del área de regiones complicadas, lo que constituye un problema básico de cálculo integral. Hay quien opina que el nuevo cálculo es pedagógicamente superior al cálculo sin infinitesimales. No obstante, ambos gozan de idéntico rigor y conducen a los mismos resultados.



3. *LOS MEDIDORES*, cuadro holandés del siglo XVII, atribuido al pintor Hendrik van Balen, ilustra las palabras del poeta latino Horacio: "En todas las cosas hay medida". No obstante, por precisas que sean las mediciones que se hagan, las cantidades infinitesimales escaparán siempre a nuestro control, puesto que a cualquier unidad de medida útil le debe corresponder un número estándar. El cuadro se expone en el Museo de Historia de la Ciencia de la Universidad de Oxford.

dos a los que se refiere su objeción, resulta que el comportamiento objetable que postula para las cosas móviles es, por lo menos, discutible. En el microrreino podríamos aplicar muchas descripciones de movimiento distintas de las que requieren recurrir a la serie completa de los puntos-observados; porque lo particular de su escenario provoca problemas conceptuales es por lo que no hay razón alguna para anatematizar la idea de movimiento.

Su segundo razonamiento, al pretender probar que un objeto jamás puede iniciar el movimiento, padece la misma enfermedad que el primero, y, por tanto, podemos rechazarla fundándonos en los mismos argumentos.

Hemos resuelto tres paradojas de Zenón utilizando algunos de los resultados técnicos de la teoría de conjuntos internos y el principio en virtud del cual los números no son adecuados para describir los hechos observados o significativos. Además, analizando la cuestión, podemos ir mucho más allá de la simple afirmación de que las paradojas de Zenón no excluyen el movimiento. Es decir, podemos construir una teoría del movimiento utilizando un resultado potente de la TCI. La teoría produce los mismos resultados que las herramientas del cálculo, es mucho más fácil de visualizar y no se ensaña con las objeciones de Zenón.

Un teorema demostrado en la TCI establece la existencia de un conjunto finito, llamado F, ¡que contiene todos los números estándar! El corolario que afirma que sólo hay un número finito de números estándar debería ser cierto, pero, sorprendemente, es falso. Para desarrollar la TCI, Nelson precisaba refinar la forma con la que los matemáticos normalmente forman sus objetos.

Una afirmación de la TCI es interna si no contiene la etiqueta "estándar". En otro caso, el objeto es externo. Los matemáticos acostumbran crear subconjuntos de conjuntos mayores predicando una propiedad que caracterice a cada uno de los objetos pertenecientes al subconjunto: las habituales bolas "rojas" o los enteros "pares". En la TCI está prohibido utilizar predicados externos, como estándar, para definir subconjuntos; la restrición se impone para evitar contradicciones. Imaginemos, por ejemplo, el conjunto de todos los números estándar de F. Este conjunto debería ser finito, puesto que es el subconjunto de un conjunto finito; y tener, por ende, un elemento mínimo, al que llamaremos r. Entonces, $r-1$ debería ser un número estándar menor que r, pero habíamos supuesto que r era el menor número estándar. Por tanto, no podemos aceptar que los números estándar son, en extensión, finitos o infinitos, dado que no podemos formar el conjunto de todos ellos y contarlos.

Con todo, el conjunto finito F, pese a su compromiso con la representación visual, nos sirve para construir nuestra teoría del movimiento. Podemos expresar ésta a la manera de una sucesión de saltos a través de F, donde cada elemento de F representa un momento distinto. De forma arbitraria, nos ceñiremos a los elementos de F que se encuentran entre 0 y 1. El tiempo 0 indica el instante en que empezamos a rastrear un objeto móvil. El segundo instante en el que podemos intentar observar el objeto móvil es el tiempo f_1, donde f_1 es el menor elemento de F que, a su vez, es mayor que 0. Si vamos así ascendiendo por F, terminaremos por alcanzar el momento f_n, siendo f_n el mayor número menor que 1. Con el instante siguiente estaremos en 1 que, en este ejemplo, es el punto de destino. Para recorrer una distancia no infinitésima (así la que separa el 0 del 1 mediante pasos infinitésimos), el subíndice n de f_n debe ser un entero ilimitado. El proceso del movimiento se divide así en $n+1$

acciones y, puesto que $n+1$ es también finito, el número de acciones puede operarse secuencialmente.

De los posibles tiempos observados, identificados con anterioridad, el progreso del objeto sólo puede acreditarse en los instantes que corresponden a ciertos números estándar de F. (En la construcción, f_1 y f_n serían no-estándar, ya que están infinitamente próximos a 0 y 1, respectivamente.) En efecto, podemos expresar un número estándar por medio de un número finito (pero no ilimitado) de cifras decimales y utilizar esta aproximación como una cierta etiqueta de su medida, pero no podemos acceder a la cola ilimitada de la expansión para alterarle un dígito y definir así un número no-estándar infinitamente próximo de su vecindad. Sólo los números estándar concretos sirven de etiquetas efectivas de medida; por lo que a la medición se refiere, la utilidad de sus entornos no-estándar es ilusoria.

Mucho hay de ilusorio en esta teoría del movimiento, y mucho queda sin expresar. Pese a todo es suficiente, en el sentido de que se puede trasladar a la notación simbólica del cálculo diferencial e integral, al que se acude para describir el movimiento. Mayor interés reviste, en este contexto, la finitud del conjunto F, que nos permite obviar las trampas de las dos primeras paradojas de Zenón. Su tercera paradoja se elude de la forma anterior: el movimiento en el tiempo real es un proceso desconocido que acontece en intervalos infinitesimales entre los puntos estándar de F; los puntos no-estándar de F, imposibles de observar, no hacen al caso.

Durante siglos, la lógica de Zenón se ha mantenido inexpugnable, prueba de la naturaleza refractaria de sus razonamientos. Pero la TCI ha logrado abrir un boquete merced a dos propiedades básicas de dicha teoría conjuntista: la capacidad para dividir un intervalo de tiempo o de espacio en un número finito de infinitesimales inefables y el carácter de los puntos estándar —los únicos que pueden ser utilizados para medir—, que son objetos discretos de la recta real. ¿Constituye nuestro trabajo sólo una solución de un viejo rompecabezas? Quizás haya varias direcciones hacia las que pudiera extenderse.

Amén de su propio valor matemático, la TCI madura con aportaciones epistemológicas, según nos acaba de mostrar su análisis. Podría modificarse para constituir una lógica epistémica general. Asimismo, los infinitesimales, o su generalización, podrían aportar un recurso técnico a la noción que Whitehead había llamado entidades reales, y que constituían los átomos generadores de su sistema filosófico. Por último, la teoría actual del movimiento y las predicciones de la física cuántica no son muy distintas, ya que ambas restringen la observación de ciertos sucesos a valores discretos. No se trata de que esta teoría del movimiento sea ninguna versión de la mecánica cuántica (ni de la teoría de la relatividad, por otra parte). Puesto que la teoría resulta de un experimento ideal, en términos de Zenón, no encierra vinculación alguna con la teoría física actual. Además, las leyes específicas inherentes a la TCI no son probablemente las más idóneas para describir la realidad. Los físicos modernos podrían adaptar la TCI modificando su sistema de leyes e introduciendo "constantes físicas", asignando quizá parámetros al conjunto F.

Pero, quizá no. Con todo, la simplicidad y elegancia de los experimentos ideales han catalizado las investigaciones a través de los siglos. Entre los ejemplos notables cabe recordar a Heinrich W. M. Olbers, cuando se pregunta por qué el cielo es obscuro de noche a pesar de la existencia de estrellas en todas las direcciones, o James Clerk Maxwell, que exige la presencia de un demonio entrometido y microscópico para criticar la segunda ley de la termodinámica. En esa onda, la argumentación de Zenón ha estimulado la reflexión sobre nuestras ideas en torno al movimiento, el tiempo y el espacio. El camino hasta su resolución se ha mostrado lleno de incidentes.

BIBLIOGRAFIA COMPLEMENTARIA

A HISTORY OF GREEK PHILOSOPHY, vol. 2: THE PRESOCRATIC TRADITION FROM PARMENIDES TO DEMOCRITUS. W. K. Guthrie. Cambridge University Press, 1965. Existe traducción española por la Ed. Gredos, Madrid.

ZENO OF ELEA. Gregory Vlastos in *The Encyclopedia of Philosophy*. Dirigido por Paul Edwards. Macmillan Publishing Company, 1967.

NONSTANDARD ANALYSIS. Martin Davis y Reuben Hersh in *Scientifican American*, vol. 226, n.º 6, págs. 78-86; junio de 1972.

INTERNAL SET THEORY: A NEW APPROACH TO NONSTANDARD ANALYSIS. Edward Nelson en *Bulletin of the American Mathematical Society*, vol. 83, n.º 6, páginas 1165-1198; noviembre de 1977.

A EPISTEMOLOGICAL USE OF NONSTANDARD ANALYSIS TO ANSWER ZENO'S OBJECTIONS AGAINST MOTION. William I. McLaughlin y Sylvia L. Miller en *Synthese*, vol. 92, n.º 3, págs. 371-384; septiembre de 1992.

La bilirrubina

Considerada antaño mero producto de desecho, el conocimiento del metabolismo de esta molécula resulta decisivo para el tratamiento de las enfermedades caracterizadas por la presencia de ictericia

Javier González-Gallego y Claudio Tiribelli

La aparición de ictericia, una coloración amarillenta de piel y mucosas, nos indica la existencia de alteraciones en la sangre y en el hígado. El fenómeno es consecuencia de un incremento en la concentración sanguínea de un pigmento biliar, la bilirrubina.

Diariamente se realizan en los hospitales de todo el mundo múltiples determinaciones de la bilirrubina. Dicho test constituye una herramienta útil para el diagnóstico y evaluación pronóstica de las enfermedades del hígado. De ahí el interés que encierra conocer mejor la estructura del pigmento y los mecanismos responsables de su metabolismo y transporte hepático.

Formada a partir de la hemoglobina y de otras proteínas que contienen la molécula hemo, la bilirrubina se transporta en sangre unida a la albúmina. Llega a las células hepáticas, en cuyo retículo endoplasmático se une a diversos azúcares. Las moléculas resultantes, los conjugados, se excretan a la bilis y se eliminan posteriormente, una vez transformados en urobilinógenos.

La concentración elevada de bilirubina en la bilis puede originar una precipitación de bilirrubinato cálcico, con la formación consiguiente de cálculos biliares. Por su parte, el incremento de los niveles de bilirrubina en sangre, denominado hiperbilirubinemia, provoca efectos tóxicos en el sistema nervioso del recién nacido, con desarrollo de encefalopatía neonatal.

Por esos y otros efectos negativos, se consideró tiempo atrás la bilirrubina un producto catabólico de desecho. Pero en 1987 un equipo de la Universidad de California en Berkeley demostró que la presencia de bilirrubina, en concentración micromolar, inhibía la oxidación inducida por radicales libres. La investigación ulterior puso de manifiesto que la bilirrubina, su precursor (la biliverdina) y algunos de los análogos de aquélla ejercían un efecto protector de las membranas celulares.

Los sistemas de defensa contra los metabolitos tóxicos abarcan enzimas y pequeñas moléculas no enzimáticas capaces de neutralizar los radicales libres, implicados en la patología del cáncer, el envejecimiento, la isquemia o la inflamación. De acuerdo con diversas investigaciones, la bilirrubina circulante podría actuar de antioxidante extracelular, protegiendo los tejidos contra los efectos deletéreos de los radicales libres.

Aislada de la bilis en 1864, hubo que esperar hasta 1942 para conocer la estructura de la bilirrubina. Su forma, o isómero, más abundante en el organismo es la bilirrubina IXα. Muestra una baja solubilidad en soluciones acuosas; ello se debe a la presencia de seis enlaces intramoleculares mediante puentes de hidrógeno, que le dan una configuración biplanar ("en teja"). Gracias a su unión con la albúmina en la circulación, se solubiliza en el plasma sanguíneo. Y se evita la formación de los puentes de hidrógeno si la bilirrubina llega, a la bilis, conjugada con diversos azúcares.

Cuando se expone el pigmento a la luz, se rompen los puentes de hidrógeno. De la disgregación resulta la fotobilirrubina y, mediante un proceso más lento, la ciclobilirrubina o lumirubina. Estos isómeros, mucho más solubles que la bilirrubina, pueden excretarse por bilis u orina. Por esa razón se utiliza la fototerapia en el tratamiento de la ictericia neonatal.

La bilirrubina proviene, en un 75 % aproximadamente, de la degradación de la hemoglobina que hay en el hígado. Los hematíes senescentes son retirados de la circulación enteros o fragmentados por lisis intravascular, con liberación de hemoglobina; los eritrocitos intactos son fagocitados por células del sistema reticuloendotelial en el hígado o el bazo. La hemoglobina circulante es captada directamente por las células hepáticas.

El 20-25 por ciento de la bilirrubina restante procede de otras proteínas que contienen el grupo hemo: citocromos y catalasa. Se llegó a ese descubrimiento a través del estudio de la incorporación, en moléculas de bilirrubina, de precursores marcados del grupo hemo (glicocola y ácido δ-aminolevulínico). Tras la inyección de glicocola radiactiva aparecen dos picos de bilirrubina radiactiva en bilis y suero. El primero, o "temprano", se observa al cabo de 72 horas. El segundo, o "tardío", se corresponde con la vida media de los eritrocitos (aproximadamente 120 días en el hombre y 50-60 días en la rata).

En la "bilirrubina temprana" distinguimos dos componentes; uno, inicial, deriva fundamentalmente de la degradación de hemoproteínas hepáticas, tales como el citocromo P-450, y aparece en bilis al cabo de unos 15 minutos tras la administración del precursor; el otro, componente último, proviene de la degradación del hemo de médula ósea, degradación que obedece a procesos de eritropoyesis ineficaz, es decir, cuando los eritrocitos no alcanzan la maduración.

A partir de los trabajos realizados a finales de los años sesenta y principios de los setenta por el grupo de Raimo Tenhunen y Rudi Schmid, de la Universidad de California, sabe-

JAVIER GONZALEZ-GALLEGO y CLAUDIO TIRIBELLI comparten investigación en los procesos hepáticos. Experto en reacciones de biotransformación hepática y en mecanismos de la secreción biliar, González-Gallego está al frente del departamento de fisiología, farmacología y toxicología de la Universidad de León. Tiribelli, que dirige el Centro de Investigación sobre el Hígado de la Universidad de Trieste, viene dedicándose desde hace años al estudio de los procesos de captación hepática.

mos que la formación de la bilirrubina implica dos pasos sucesivos. La molécula hemo se convierte primero en biliverdina por medio de una hemooxigenasa microsómica. Esta enzima se localiza en el retículo endoplasmático de hepatocitos, de células de Kupffer, de células del bazo y de otros tejidos; la enzima muestra una elevada especificidad hacia el enlace en posición α del hemo, lo que explica que la mayoría de la biliverdina y bilirrubina formadas correspondan a isómeros IXα.

La reacción se bloquea por diversas metalporfirinas. Y en éstas se piensa a la hora de prevenir la hiperbilirrubinemia neonatal o el síndrome de Crigler-Najjar. En 1988, el equipo de Atallah Kappas, de la Universidad Rockefeller, acometió un ensayo de doble ciego que involucró a un total de 517 recién nacidos prematuros. Comprobó que, en los bebés que recibieron una única y pequeña dosis de Sn-protoporfirina después del nacimiento, los niveles de bilirrubina sérica se redujeron en un 41 % y la necesidad de fototerapia se eliminó en un 76 %. La razón estriba en que esa metalporfirina bloquea la acción de la hemooxigenasa microsómica. El grupo de Kappas investiga ahora los efectos de la Zn-protoporfirina, que presenta un menor riesgo de fotosensibilización y varias ventajas; entre éstas, su mayor dificultad para atravesar la barrera hematoencefálica y la posibilidad de administración oral, lo que facilita la distribución hepática del fármaco.

En aves, reptiles y anfibios, el precursor biliverdina se excreta directamente a la bilis y constituye el producto final de la degradación del hemo. En mamíferos, la biliverdina se transforma en bilirrubina por la biliverdina reductasa citoplasmática, enzima que muestra una elevada actividad en hígado, bazo y riñón.

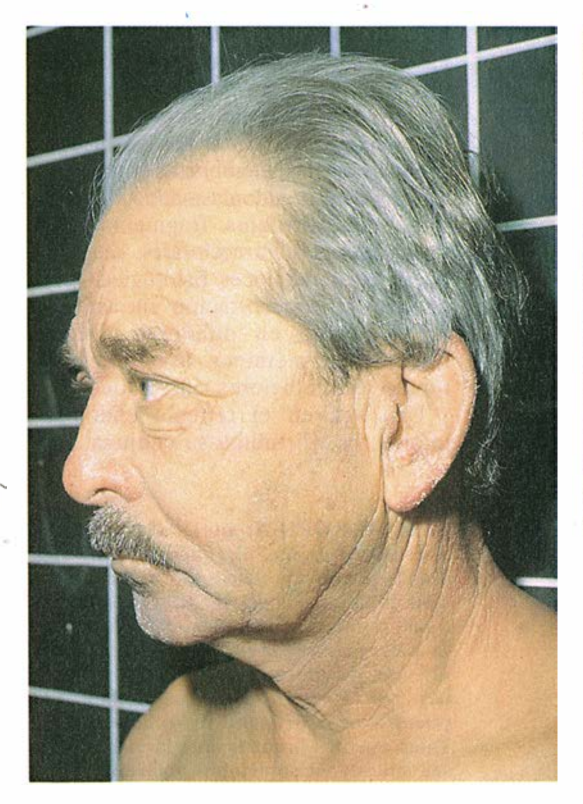
No es fácil entender a primera vista la reducción de biliverdina a bilirrubina y su posterior conjugación. Por un lado, se trata de reacciones costosas desde el punto de vista energético y, por otro, la bilirrubina resulta tóxica cuando se acumula en concentraciones elevadas. De acuerdo con la hipótesis dominante, la formación de bilirrubina sería necesaria para su eficaz transferencia del feto a la madre a través de la placenta, pues cruza mejor las membranas biológicas que la biliverdina.

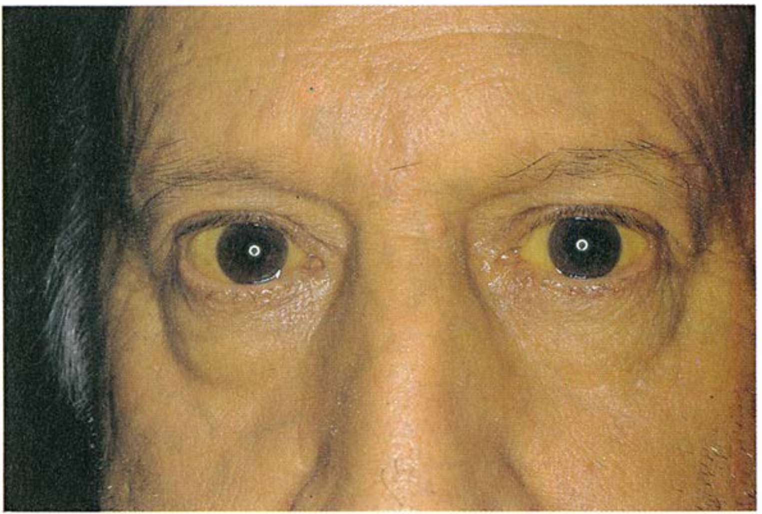
Algunos mamíferos, sin embargo, excretan también biliverdina. En 1983, uno de nosotros (González) demostró que en el conejo se excretan bilirrubina y biliverdina en concentraciones muy similares. En un estudio posterior comprobó que el fenómeno se debía al carácter limitante de la baja actividad de la enzima biliverdina reductasa. La presencia de ambos pigmentos biliares en la bilis del conejo fue confirmada por el grupo de Charles Cornelius, de la Universidad de California en Davis, que descubrió la excreción de biliverdina en otro mamífero, la nutria.

Además, la formación de bilirrubina no es exclusiva de animales placentarios. Muchos peces excretan este pigmento biliar. Quizá cuando se confirme el papel de la bilirrubina como antioxidante endógeno tengamos por fin una razón que nos explique su aparición en el curso de la evolución.

Una vez formada, la bilirrubina se transporta en sangre unida a albúmina. A esa unión se oponen los ácidos grasos libres, las sulfonamidas, los salicilatos y la ampicilina, entre otros; al impedir dicha fijación, se produce el desplazamiento de la bilirrubina y el aumento de ésta libre, que, así, puede atravesar la barrera hematoencefálica y provocar lesiones del sistema nervioso. Sin embargo, se requieren concentraciones de bilirrubina muy elevadas o situaciones de hipoalbuminemia para que el desplazamiento tenga lugar.

En pacientes con hiperbilirrubinemia conjugada prolongada, en los que los niveles sanguíneos de los conjugados de bilirrubina formados en el hígado son muy altos, parte de la bilirrubina se une irreversiblemente a la albúmina. Dicha fracción, denominada δ-bilirrubina o bil-alb, puede llegar a suponer un porcentaje importante de la bilirrubina plasmática total en las enfermedades hepáticas. No se elimina libremente por el ri-





1. LA ICTERICIA CONSTITUYE la traducción clínica de la acumulación de bilirrubina en el organismo. Los sujetos con ictericia desarrollan una coloración amarillenta de la piel y de las mucosas debida al depósito de la bilirrubina. Obsérvese en la fotografía de la derecha cómo dicha coloración es especialmente patente en la esclerótica.

ñón, y es, por tanto, responsable de la persistencia de niveles elevados de bilirrubina durante la fase de resolución de la enfermedad.

El hígado es el órgano encargado del aclaramiento de la bilirrubina procedente del plasma. Antes de que tenga lugar la captación, el pigmento se disocia de la albúmina y atraviesa entonces la membrana sinusoidal de los hepatocitos, es decir, la situada en el polo celular próximo a los sinusoides o capilares sanguíneos del hígado.

En la membrana sinusoidal existen al menos tres proteínas responsables de la captación de la bilirrubina y de diversos aniones orgánicos. La primera es la bilitraslocasa (BTL), una proteína con un peso molecular de 37 kilodalton descubierta por uno de nosotros (Tiribelli), Gian Luigi Sottocassa y colaboradores en la Universidad de Trieste. La utilización de anticuerpos específicos frente a esta proteína inhibe la captación de bromosulfoftaleína y otros aniones orgánicos, pero no de los ácidos biliares.

Un segundo sistema de transporte se relaciona con una proteína con un peso molecular de 54-55 kilodalton, descubierta por Allen W. Wolkoff, de la Universidad de Nueva York, a la que se ha denominado proteína fijadora de aniones orgánicos (OABP). Los anticuerpos frente a dicha proteína inhiben la captación de bilirrubina y bromosulfoftaleína en hepatocitos aislados a bajas concentraciones y, como en el caso anterior, no se altera el transporte de los ácidos biliares.

Finalmente, Wolfgang Stremmel, de la Universidad de Düsseldorf, y Paul Berk, de la de Nueva York, han descrito un tercer sistema de transporte de la bilirrubina en el que está implicada una proteína con un peso molecular de 55 kilodalton. Se trata de la proteína fijadora de bromosulfoftaleína/bilirrubina (BBBP); es, probablemente, responsable del transporte de baja afinidad de la bromosulfoftaleína a concentraciones elevadas.

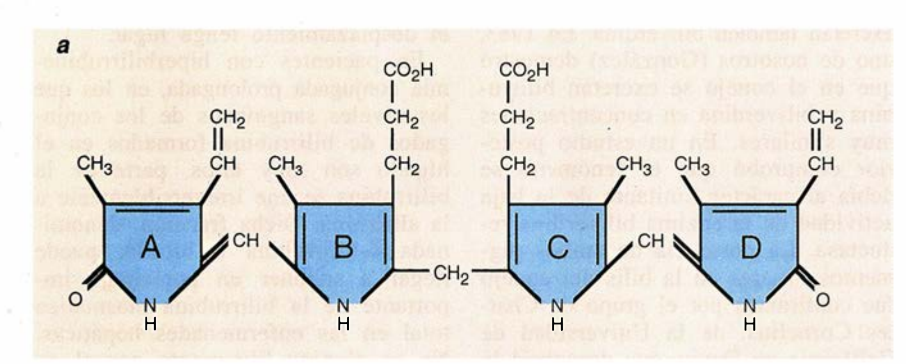
La contribución relativa de cada una de las tres proteínas en la captación hepatocitaria de aniones orgánicos, todavía por determinar, podría depender de la concentración del sustrato.

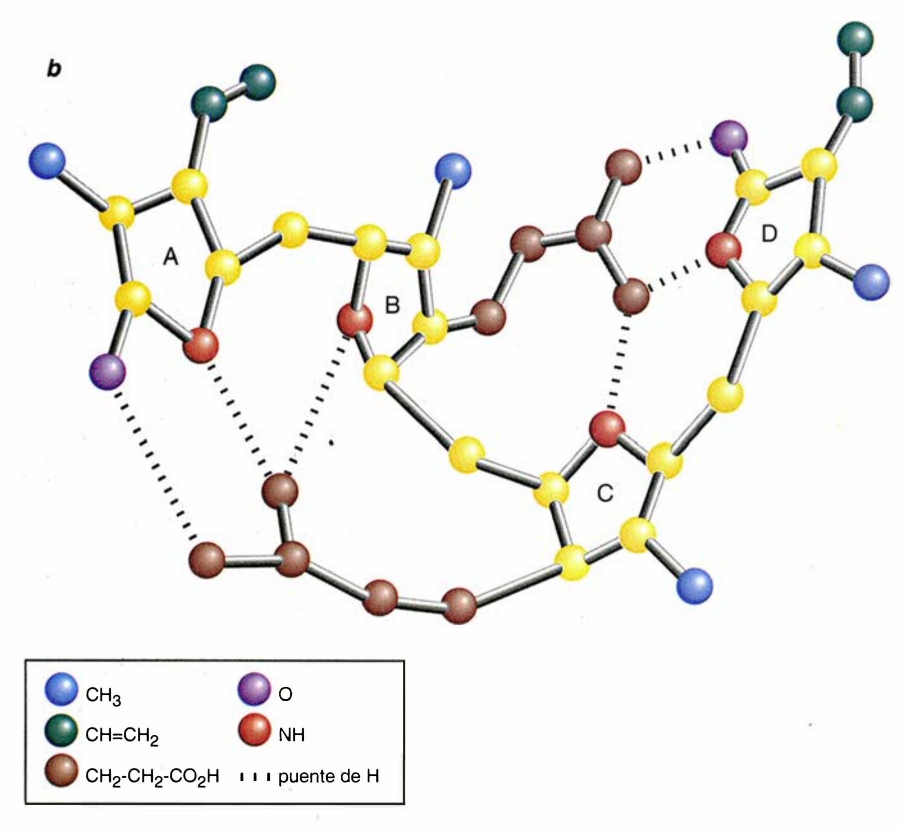
Ya tenemos la bilirrubina en el interior del hepatocito. ¿Qué transformaciones experimenta? Se une a las proteínas Y (ligandina) y Z, ambas intracelulares. De éstas, la ligandina constituye hasta un 5 % de las proteínas totales en el citoplasma de los hepatocitos; en menor proporción aparece en el riñón, intestino delgado y placenta. La función de la fijación a la ligandina podría ser la de evitar el reflujo del pigmento a la circulación y su difusión inespecífica a diferentes compartimentos de las células del hígado.

La proteína Z, que parece ser idéntica a la proteína fijadora de ácidos grasos (FABP), presenta menos afinidad por la bilirrubina. Implicada en el transporte intracelular de ácidos grasos, la proteína Z se encuentra en la mucosa intestinal, hígado, miocardio y otros tejidos; se ignora si interviene en el transporte hepático y almacenamiento intracelular de la bilirrubina.

Aunque la bilirrubina pasa de la membrana sinusoidal a las membranas del retículo endoplasmático unida a proteínas intracelulares, cabría también la posibilidad de que lo hiciera aprovechando el contacto directo entre membranas. Se ha demostrado que la bilirrubina puede incorporarse en las zonas hidrofóbicas de las bicapas lipídicas de las membranas y alcanzar así el retículo endoplasmático. La cantidad de bilirrubina fragmentada en las membranas intracelulares, muy pequeña en condiciones fisiológicas, crece con la concentración intracelular del pigmento en diferentes estados hiperbilirrubinémicos o bajo el estímulo de inductores enzimáticos.

Una vez en el retículo endoplasmático, la bilirrubina se conjuga, for-





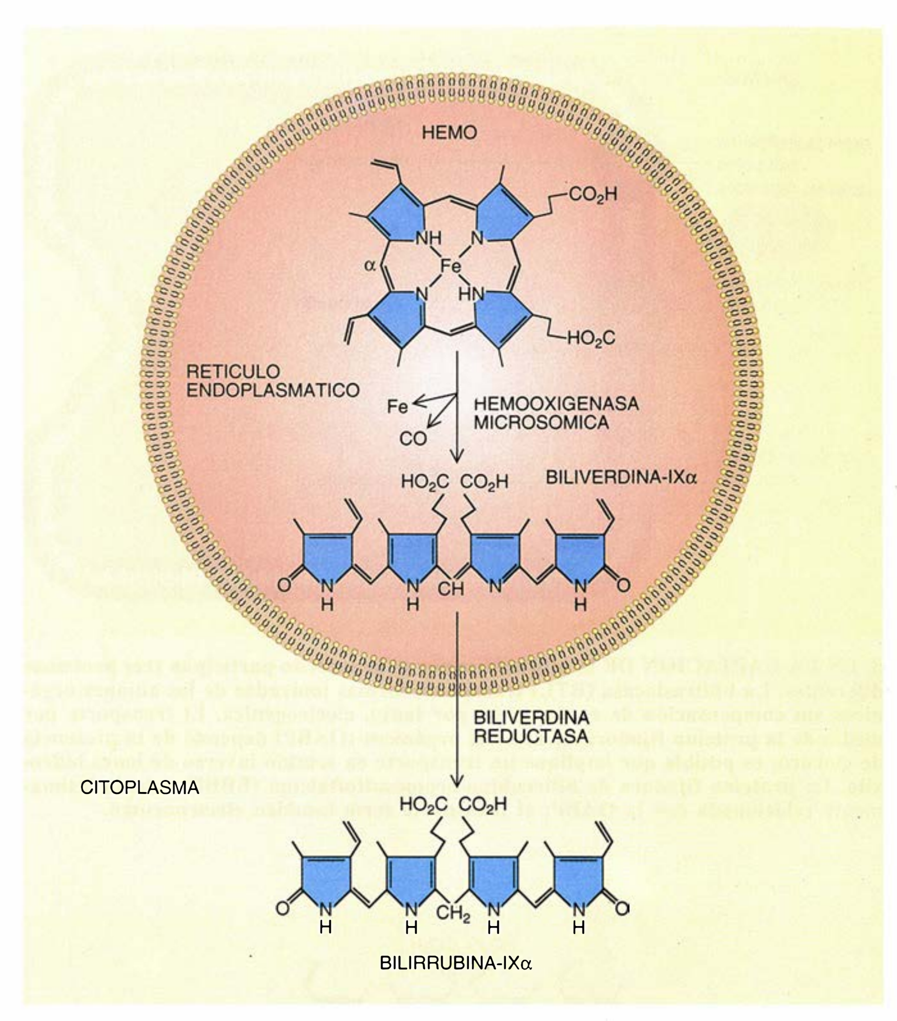
2. BILIRRUBINA, molécula formada por cuatro anillos pirrólicos unidos mediante puentes meténicos. La configuración de la parte superior (*a*) corresponde a una representación lineal de la molécula de bilirrubina IXα Z,Z. La configuración inferior (b) muestra la estructura biplanar "en teja", en la que se forman seis puentes de hidrógeno intramoleculares. Tales puentes impiden que la bilirrubina sea soluble en soluciones acuosas.

ma complejos, mediante la unión con azúcares. De dicha fijación saldrán compuestos monoconjugados y diconjugados, según que la bilirrubina esté unida a una o a dos moléculas del azúcar.

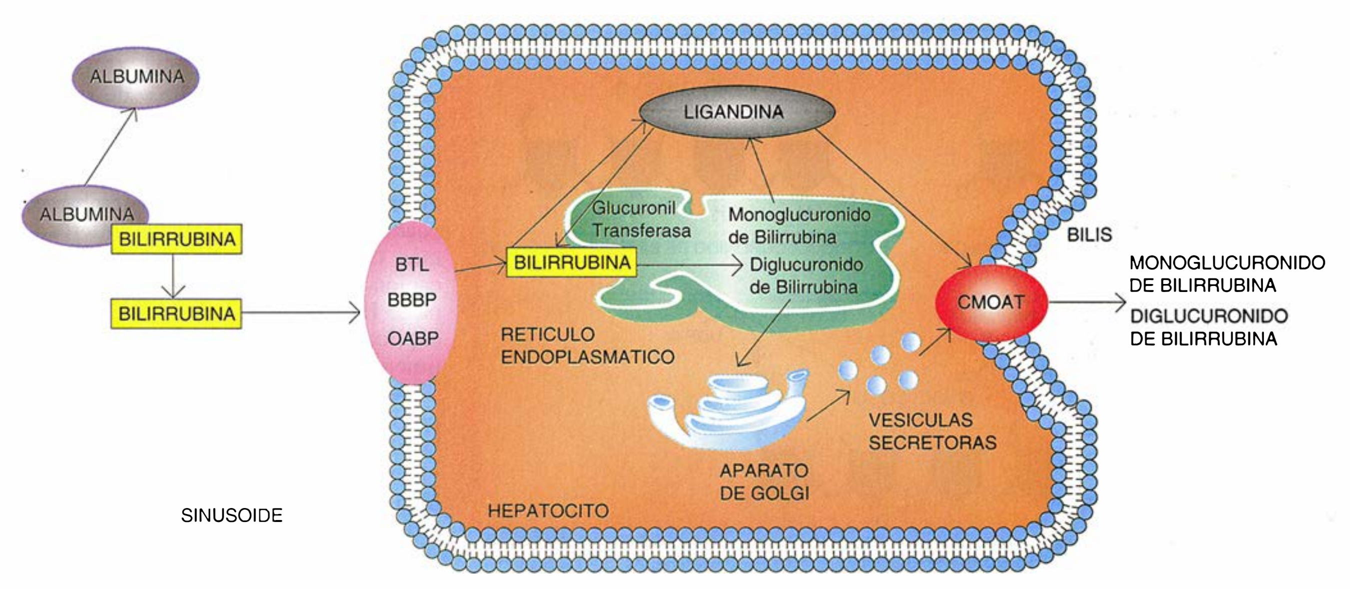
Pero, ¿de qué bilirrubina estamos ahora hablando? En 1916 se demostró la existencia de dos tipos de bilirrubina en plasma. El pigmento presente en la sangre de pacientes con enfermedades hemolíticas requería la adición de alcohol para dar una reacción positiva con el ácido sulfanílico diazotizado (reactivo de Ehrlich); a esa bilirrubina se la denominó "indirecta". En la bilis y en el plasma de pacientes con ictericia obstructiva se detectaba una bilirrubina "directa", que no requería la presencia de alcohol para dar la reacción. Y se concluyó que la molécula de bilirrubina sufría cambios a su paso a través del hígado.

Sin embargo, en 1953 se comprobó que las diferencias obedecían a la presencia de dos fracciones diferentes: la bilirrubina "directa" era una mezcla de dos pigmentos, formados en el hígado, que se denominaron I y II. Años más tarde, se confirmó que dichas fracciones se correspondían con moléculas de bilirrubina monoconjugada y diconjugada, respectivamente, mientras que la bilirrubina "indirecta" sería el pigmento sin conjugar.

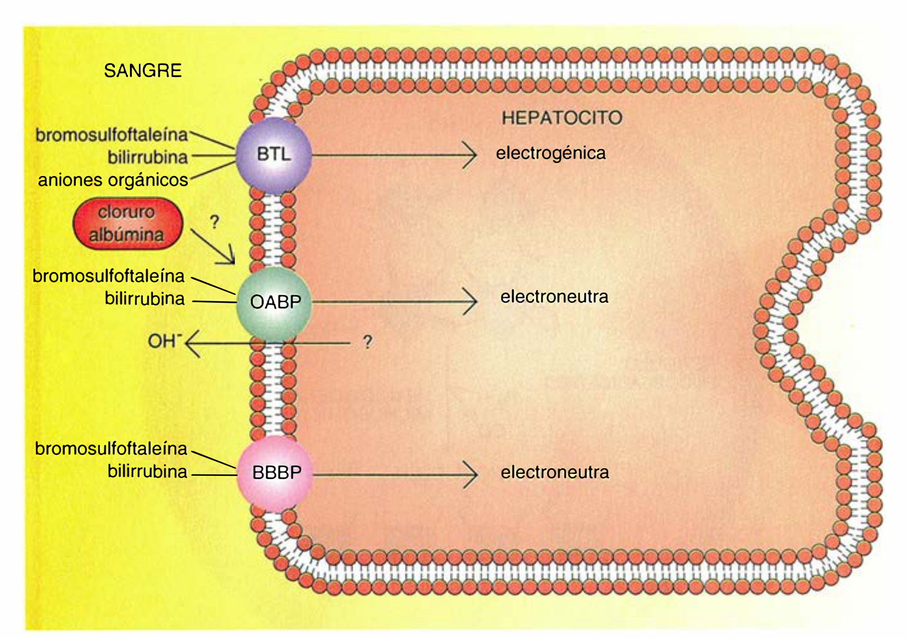
La bilirrubina puede dar lugar a la formación de dos monoconjugados, según que la conjugación se produzca en las posiciones C8 o C12. La



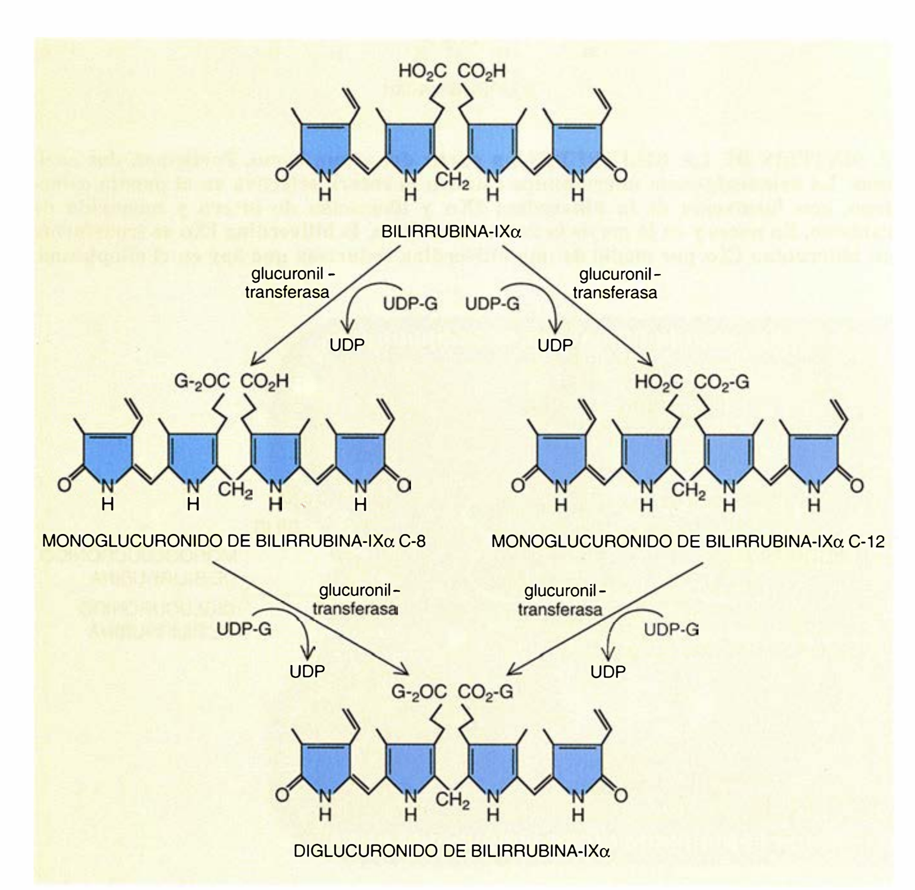
3. SINTESIS DE LA BILIRRUBINA a partir del grupo hemo. Participan dos enzimas. La hemooxigenasa microsómica cataliza la rotura selectiva en el puente α-meteno, con formación de la biliverdina IXα y liberación de hierro y monóxido de carbono. En peces y en la mayoría de los mamíferos, la biliverdina IXα se transforma en bilirrubina IXα por medio de una biliverdina reductasa que hay en el citoplasma.



4. TRANSPORTE DE BILIRRUBINA en los hepatocitos. La bilirrubina se une a la albúmina en la circulación, pero el complejo resultante se disocia para que la bilirrubina penetre en la célula, lo que hace con el concurso de proteínas transportadoras (BTL, BBBP y OABP). El pigmento se dirige al retículo endoplasmático unido a proteínas citoplasmáticas o a través del contacto directo entre membranas. La conjugación, catalizada por la UDP-glucuroniltransferasa, origina monoconjugados y diconjugados que se transportan a la membrana canalicular por difusión o por medio de la vía microtúbulo-vesicular. El sistema cMOAT se encarga finalmente de la excreción al canalículo biliar.



5. EN LA CAPTACION DE la bilirrubina por el hepatocito participan tres proteínas diferentes. La bilitraslocasa (BTL) transporta formas ionizadas de los aniones orgánicos sin compensación de cargas y es, por tanto, electrogénica. El transporte por medio de la proteína fijadora de aniones orgánicos (OABP) depende de la presencia de cloruro; es posible que implique un transporte en sentido inverso de iones hidroxilo. La proteína fijadora de bilirrubina/bromosulfoftaleína (BBBP) estaría íntimamente relacionada con la OABP; el transporte sería también electroneutro.



6. CONJUGADOS DE BILIRRUBINA. Se forman por unión a una o a dos moléculas de ácido UDP-glucurónico (UDP-G). Los monoglucurónidos de bilirrubina se sintetizan por reacción con uno de los grupos de ácido propiónico del pigmento biliar en las posiciones C8 o C12. Los monoglucurónidos formados se convierten en diglucurónido de bilirrubina. Las reacciones están catalizadas por la UDP-glucuroniltransferasa.

conjugación está catalizada por la bilirrubina UDP-glucuroniltransferasa. Los monoconjugados se transforman más tarde en bilirrubina diconjugada en una nueva reacción catalizada por la misma enzima. ¿Qué estructura tiene ésta?

En el hombre se encuentran múltiples isoenzimas, que se clasifican en dos subfamilias de acuerdo con su especificidad de sustrato y con la secuencia de los ADN complementarios. La subfamilia 1 consta de, al menos, cuatro glucuroniltransferasas, que catalizan la conjugación de fenoles y de bilirrubina y que están codificadas por genes situados en el cromosoma 2. En esta subfamilia se han identificado dos isoenzimas diferentes (HUGBr1 y HUGBr2), capacitadas para catalizar la conjugación de la bilirrubina. La subfamilia 2 consta de al menos cinco glucuroniltransferasas, codificadas por genes del cromosoma 4, que catalizan la conjugación de esteroides o ácidos biliares.

El grupo de Karel Heirwegh y Johan Fevery, de la Universidad de Lovaina, ha estudiado los tipos de conjugados de bilirrubina formados en diversas especies animales. En el hombre predominan los diglucurónidos, con cantidades pequeñas de glucósidos y xilósidos. En otros mamíferos se manifiestan pautas muy diversas, con predominio de los monoglucurónidos en el ratón, la oveja o el conejo, y predominio de diglucurónidos en la rata, el perro, el gato o el mono.

En condiciones fisiológicas, la membrana microsómica es muy poco permeable a sustratos polares. (Los microsomas son estructuras derivadas del retículo endoplasmático, rodeadas por una membrana y matriz compacta.) La actividad de la UDP-glucuroniltransferasa aumenta, sin embargo, cuando se permeabilizan los microsomas con detergentes, mediante sonicación o por tratamiento con fosfolipasa. Han aparecido dos modelos para explicar dicha activación: el compartimental y el de constricción.

De acuerdo con el primer modelo, propuesto por Norbert Blanckaert, de la Universidad de Lovaina, la UDP-glucuroniltransferasa se situaría en el lado interno de las membranas del retículo endoplasmático y habría un sistema transportador (una translocasa) que posibilitaría el acceso de los UDP-azúcares, a través de la barrera membranosa, hasta el compartimento interno; los agentes permeabilizantes permitirían la expresión de la enzima al evitar la fase de translocación.

De acuerdo con el modelo de constricción, avanzado por el grupo de David Zakim, de la Universidad de Cornell, el efecto activador de los agentes permeabilizantes sería indirecto, producido por alteraciones en la organización del microentorno de la membrana microsómica; ello induciría cambios en la conformación de la enzima. No podemos decidir todavía cuál es el modelo acertado.

En el hombre, la conjugación de la bilirrubina con ácido glucurónico sólo se desarrolla por completo tras el nacimiento. A las catorce semanas de edad se alcanzan ya los niveles del adulto. La baja actividad UDP-glucuroniltransferasa en el momento de nacer tiene que ver con la aparición de la hiperbilirrubinemia fisiológica del recién nacido. Pero intervienen también otros factores; la producción elevada de bilirrubina (la baja actividad UDP-glucuroniltransferasa) y las dificultades de perfusión sanguínea del hígado contribuyen a la ictericia.

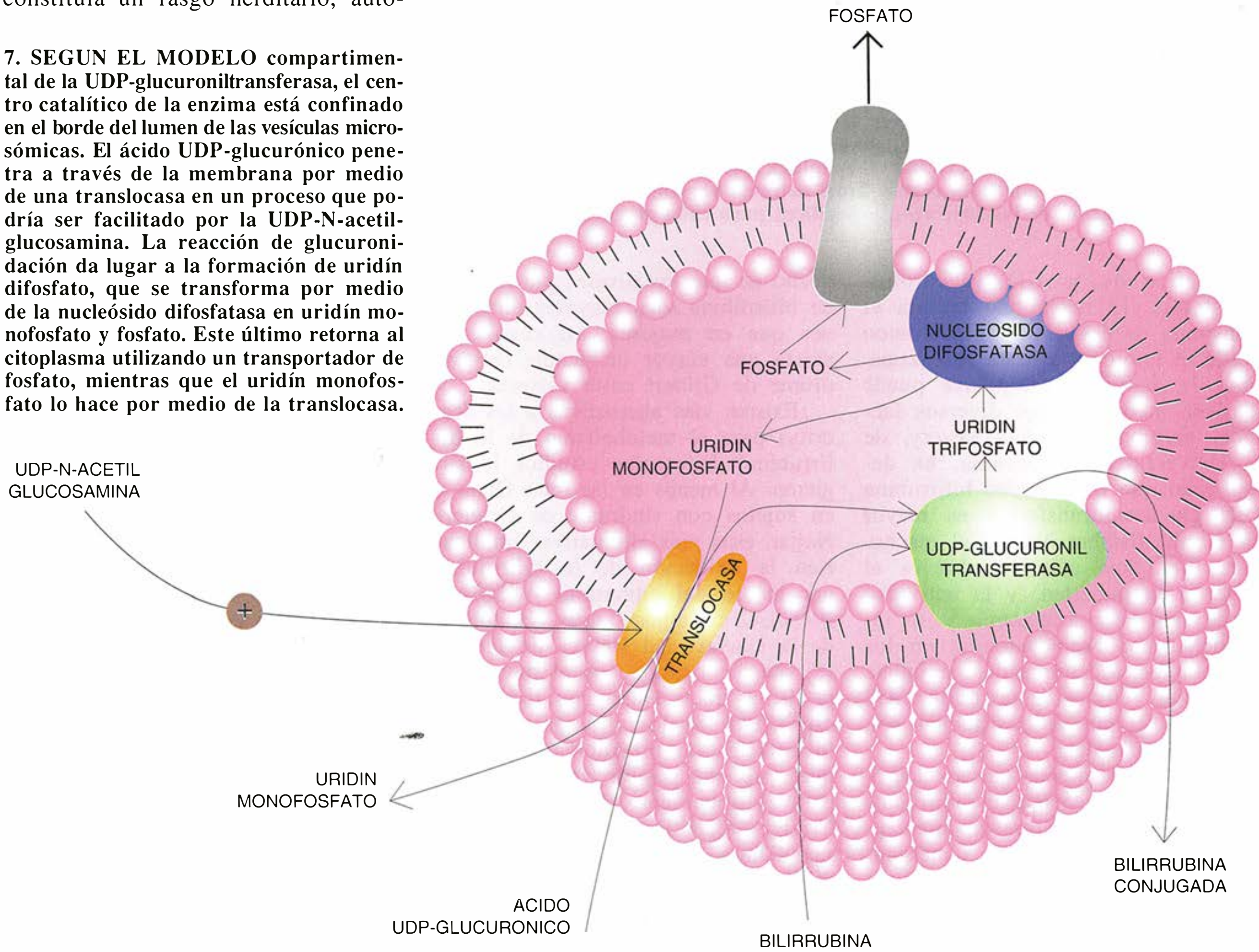
En 1939, se describió una cepa mutante de ratas, la rata Gunn, que presentaba una desmesurada concentración de bilirrubina, no conjugada, en sangre. Esa hiperbilirrubinemia constituía un rasgo herditario, autosómico y recesivo. Se comprobó más tarde que los animales no desarrollaban una hemólisis perceptible y se sugirió la posibilidad de que tuvieran problemas con la formación de bilirrubina "directa". Hoy sabemos que las ratas Gunn poseen formas defectuosas de la bilirrubina UDP-glucuroniltransferasa, hasta el punto de quedar anulada la conjugación de la bilirrubina en los homocigotos.

El síndrome de Crigler-Najjar y el síndrome de Gilbert se caracterizan por deficiencias de la actividad glucuroniltransferasa. El primero, de escasa incidencia, puede presentarse bajo dos tipos. En el síndrome de Crigler-Najjar de tipo I existe una hiperbilirrubinemia intensa, asociada con la aparición de encefalopatía y la muerte en la infancia. El tipo II muestra una hiperbilirrubinemia más suave y fluctuante; no suele desarrollar encefalopatías.

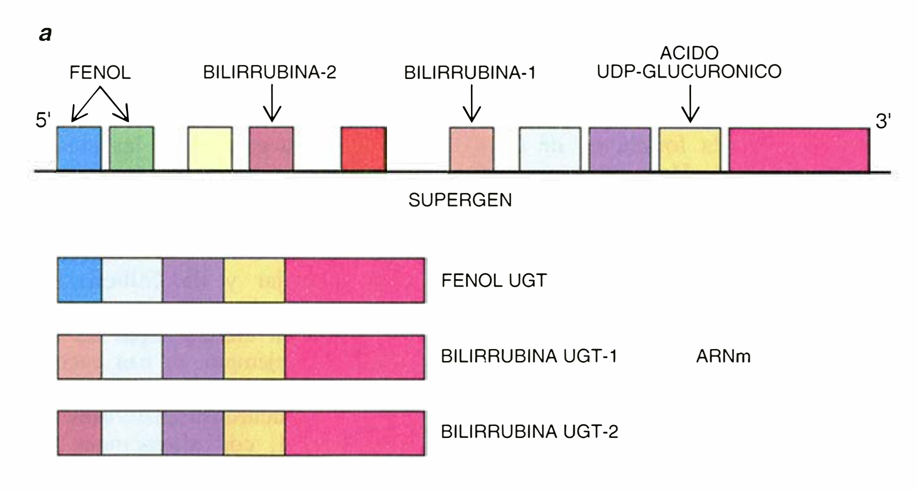
Por lo que se refiere al síndrome de Gilbert, de altísima incidencia entre la población occidental (de un 2 % a un 12 %), viene éste definido por una moderada hiperbilirrubinemia. En esta situación morbosa parecen coexistir una limitada actividad de la glucuroniltransferasa y anomalías en la captación hepática de la bilirrubina.

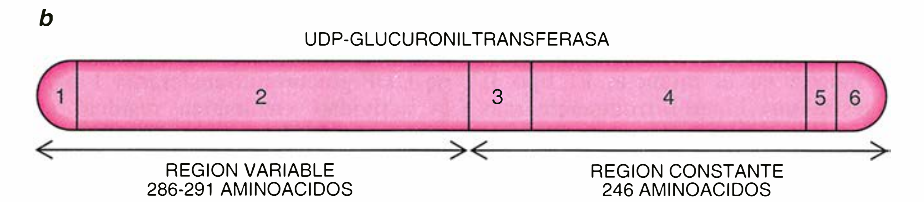
La estructura general de las glucuroniltransferasas y de los genes que las determinan explican los fallos que pueden detectarse en la actividad enzimática de pacientes con síndromes de Crigler-Najjar y de Gilbert. En algunos afectos de síndrome de Crigler-Najjar y en ciertas cepas de ratas Gunn, por ejemplo, se han encontrado lesiones que producen anomalías de todas las glucuroniltransferasas en la subfamilia 1, con alteraciones en la conjugación de fenoles y de bilirrubina. En pacientes con el síndrome de Crigler-Najjar de tipo II o con el síndrome de Gilbert, parece haber anomalías selectivas de las bilirrubina UDP-glucuroniltransferasas 1 o 2; la actividad enzimática residual se debería a la isoenzima no alterada. Los defectos incompletos en la actividad conjugante pueden corregirse, al menos parcialmente, mediante tratamiento con fenobarbital.

Según parece, el nucleótido cíclico UDP-N-acetilglucosamina desempeña un papel regulador en la actividad



7. SEGUN EL MODELO compartimental de la UDP-glucuroniltransferasa, el centro catalítico de la enzima está confinado en el borde del lumen de las vesículas microsómicas. El ácido UDP-glucurónico penetra a través de la membrana por medio de una translocasa en un proceso que podría ser facilitado por la UDP-N-acetilglucosamina. La reacción de glucuronidación da lugar a la formación de uridín difosfato, que se transforma por medio de la nucleósido difosfatasa en uridín monofosfato y fosfato. Este último retorna al citoplasma utilizando un transportador de fosfato, mientras que el uridín monofosfato lo hace por medio de la translocasa.





1. SECUENCIA DE SEÑAL
2. UNION DEL SUSTRATO
3. SITIOS DE GLICOSILACION
4. UNION DEL ACIDO UDP-GLUCURONICO
5. TRANSMEMBRANA
6. DETENCION DE TRANSFERENCIA

8. EN EL SUPERGEN de la familia 1 de las UDP-glucuroniltransferasas (*a*), la especificidad de sustrato de cada enzima viene determinada por la separación, en el proceso de traducción, de un exón que determina la región de unión al sustrato. Llamamos exones a los segmentos de ADN que se transcriben, es decir, que dan lugar a la producción de ARNm. En el esquema, cada recuadro representa un exón y las líneas corresponden a los intrones, los segmentos que no se transcriben. Todos los polipéptidos de la familia 1 de UDP-glucuroniltransferasas (*b*) presentan la misma secuencia aminoacídica en el extremo C-terminal, que incluye la región de fijación del ácido UDP-glucurónico. La región 2 del extremo variable N-terminal contiene los sitios de fijación de los sustratos, específicos para cada glucuroniltransferasa.

enzimática de la UDP-glucuroniltransferasa. El nucleótido facilita el transporte de ácido UDP-glucurónico desde el citoplasma hasta la enzima. Pero la actividad enzimática puede también modularse por diversos factores hormonales. Johan Fevery, de la Universidad de Lovaina, ha demostrado que la actividad bilirrubina UDP-glucuroniltransferasa es mayor en la rata hembra que en el macho. Esta diferencia sólo se observa al llegar a la pubertad, y la actividad decae tras la ovariectomía y se incrementa por orquidectomía; en los animales gonadectomizados, se puede lograr un retorno a la normalidad mediante terapia de sustitución con hormonas sexuales.

De ello se desprende que los andrógenos reducen la actividad bilirrubina UDP-glucuroniltransferasa, mientras que las hormonas sexuales femeninas inducen el efecto opuesto. Dichas diferencias, extrapoladas al ser humano, podrían contribuir a explicar el hecho de que los niveles plasmáticos de bilirrubina sean mayores en varones que en mujeres, así como que exista una mayor incidencia de síndrome de Gilbert entre varones.

¿Existen vías alternativas a las descritas para el metabolismo de la bilirrubina? Diferentes estudios lo sugieren. Al menos en las ratas Gunn y en sujetos con síndrome de Crigler-Najjar, esas vías alternativas implicarían la oxidación de la bilirrubina por medio del citocromo P-450IA1, P-450IA2 o ambos. Jaime Kapitulnik, de la Universidad Hebrea de Jerusalén, ha demostrado que el tratamiento de las ratas Gunn con inductores específicos de los citocromos rebaja los niveles de bilirrubina sérica, lo que abre una posible vía terapéutica para el tratamiento del síndrome de Crigler-Najjar. De hecho, se está desarrollando un estudio acerca de la eficacia del tratamiento a largo plazo con indol-3-carbinol en pacientes con síndrome de Crigler-Najjar tipo I.

Aunque hace más de treinta años que conocemos la existencia de los monoconjugados y diconjugados de bilirrubina, hasta principios de los ochenta no se pudo disponer de métodos para su determinación, tanto en sangre como en bilis, y alcanzar así un conocimiento más profundo del metabolismo de la bilirrubina. Entre ellos, destaca la cromatografía líquida de alta eficacia (HPLC).

Gracias a la HPLC podemos distinguir tres grupos distintos de alteraciones en el metabolismo de la bilirrubina. En la hiperbilirrubinemia no conjugada pura, característica de los síndromes de Gilbert y Crigler-Najjar, aparece un incremento selectivo en la concentración sérica de bilirrubina no conjugada, con concentraciones de bilirrubina monoconjugada y diconjugada normales o reducidas y un porcentaje de conjugados respecto al pigmento total disminuido.

La hiperbilirrubinemia predominantemente conjugada, que aparece en las enfermedades hepatocelulares o en la obstrucción extrahepática, se caracteriza por grandes incrementos en la concentración de conjugados y valores de bilirrubina no conjugada normales o ligeramente elevados; en consecuencia, se eleva el porcentaje de conjugados.

Por último, en situaciones en las que existe una hiperproducción de bilirrubina, tales como la hemólisis o la eritropoyesis ineficaz, se dan incrementos paralelos en la concentración de la bilirrubina no conjugada y de los conjugados, manteniéndose el porcentaje de éstos dentro de valores normales.

Entramos, por fin, en la postrera etapa, el tránsito de la bilirrubina desde el hígado hacia los conductos biliares. Los monoconjugados y diconjugados de la bilirrubina se segregan a los canalículos biliares por un proceso dependiente de energía; el pigmento se transporta contra gradientes de concentración. Ese proceso opera a una capacidad menor que el proceso de captación. La fracción de bilirrubina no conjugada que podamos encontrar en la bilis deriva probablemente de pigmento monoconjugado en una reacción catalizada por una β-glucuronidasa. Dicha enzima, presente en lisosomas y en microsomas, parece modular la tasa de glucuronidación de la bilirrubina; se ha observado que los ratones mutantes con deficiencia de β-glucuronidasa excretan mayores cantidades de bilirrubina conjugada, tanto en situaciones basales como de hiperbilirrubinemia.

El sistema de transporte canalicu-

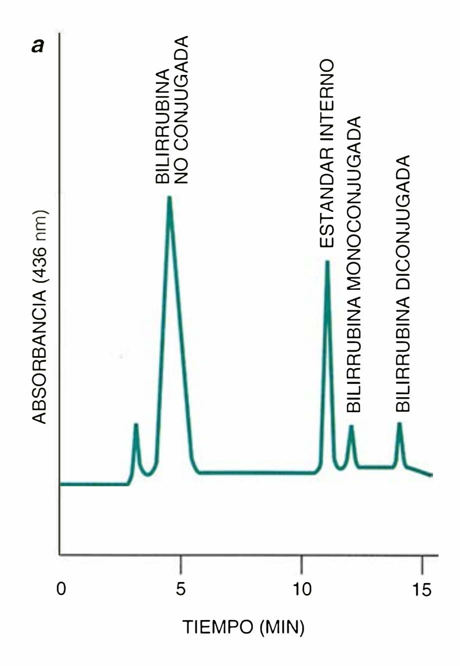
lar de la bilirrubina lo comparten diferentes aniones orgánicos, pero no los ácidos biliares. Tenemos así que, en los síndromes de Dubin Johnson y Rotor en el hombre y de la oveja mutante Corriedale, caracterizados por hiperbilirrubinemia y anomalías en la secreción de bilirrubina y otros aniones orgánicos, la secreción y los niveles séricos de ácidos biliares se mantienen en niveles normales.

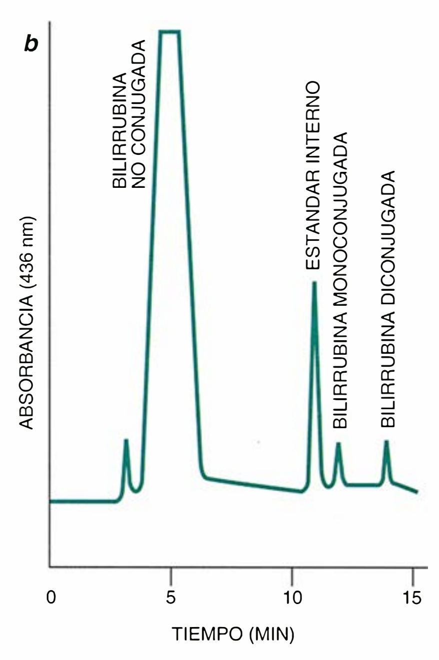
En 1983 Wout Lamers, de la Universidad de Amsterdam, descubrió una cepa de ratas mutantes que mostraban una marcada hiperbilirrubinemia, correspondiente en su mayor parte a monoglucurónidos y diglucurónidos de bilirrubina. Cuatro años después se detectaron también en la Universidad de Groningen y en Japón animales con hiperbilirrubinemia conjugada. Se ha comprobado que todas las cepas manifiestan un mismo defecto autosómico recesivo. Los estudios llevados a cabo por el equipo de Peter Jansen, de la Universidad de Amsterdam, revelan que los animales mutantes tienen una secreción normal de ácidos biliares, pero existe una notable reducción en la de otros aniones: bilirrubina, bromosulfoftaleína, cisteinil leucotrienos o glutatión oxidado.

Una misma proteína transporta todos estos compuestos. La cMOAT (del inglés c*analicular* Multi*O*rg*anic* A*nion* T*ransporter*) constituye, al parecer, una proteína integral de la membrana plasmática canalicular del hepatocito, que obtiene la energía para el transporte por hidrólisis del ATP. Dicha proteína sería responsable de la excreción canalicular de aniones orgánicos divalentes.

A pesar de utilizar transportadores canaliculares diferentes, los ácidos biliares pueden modificar la excreción biliar tanto de la bilirrubina como de otros aniones orgánicos. La secreción biliar máxima de la bilirrubina aumenta tras la infusión de diferentes ácidos biliares, aunque los resultados obtenidos varían bastante en función de la dosis y de la especie estudiada. Para explicar ese efecto estimulante se postuló una incorporación de la bilirrubina a las micelas mixtas formadas por los ácidos biliares en la bilis. Pero no sólo entonces. Como ha demostrado uno de los autores (González), el efecto se presenta también tras la infusión de ácidos biliares sintéticos no formadores de micelas.

Tras la infusión de ácidos biliares se produce el reclutamiento de un sistema que se apoya en los microtúbulos y que es capaz de incrementar el transporte de la bilirrubina a través de los hepatocitos. James M. Crawford y John L. Gollan, de la Universidad de Harvard, han propuesto un modelo según el cual, en condiciones basales, la mayoría de los conjugados de bilirrubina se unirían a las proteínas citoplasmáticas, para ser transportados, por difusión, hasta la membrana canalicular. Ahora bien, al aumentar el flujo de ácidos biliares, una parte de éstos y de la bilirrubina conjugada podría transportarse a través de microtúbulos y vesículas. De hecho, estos investigadores han comprobado que la colchicina, un agente que altera la estructura de los microtúbulos, no influye en la secreción de bilirrubina a menos que las concentraciones intracelulares de ácidos biliares se eleven sobre su nivel basal.





9. CROMATOGRAMAS obtenidos por técnicas de cromatografía líquida de alta eficacia en sueros procedentes de un sujeto normal (*a*) y de un paciente con el síndrome de Gilbert (*b*). En el primer caso el porcentaje de bilirrubina monoconjugada más diconjugada supone un 3,5 % de la bilirrubina sérica total. En el segundo caso hay una elevación de la concentración de bilirrubina no conjugada, no se modifica significativamente la concentración de conjugados y se reduce el porcentaje de éstos a un 0,9 %.

La bilirrubina conjugada no se absorbe en el intestino ni en la vesícula biliar. La degradan las bacterias colónicas en urobilinógenos, un conjunto de tetrapirroles. Aunque los urobilinógenos pueden presentar una circulación enterohepática, sólo una pequeña fracción se reabsorbe desde el intestino y más del noventa por ciento de los mismos se reexcreta de nuevo en la bilis. En neonatos o en casos de parálisis intestinal, puede producirse desconjugación de la bilirrubina; si así ocurre, el pigmento absorbido fomenta la formación de una hiperbilirrubinemia no conjugada.

En los últimos años se ha registrado un considerable progreso en nuestro conocimiento de las diferentes fases del transporte hepatocelular y del metabolismo de la bilirrubina. Los hallazgos realizados por diferentes grupos ponen de manifiesto la complejidad de los procesos fisiológicos implicados y los múltiples componentes requeridos. El reto que se nos plantea es llegar a definir los detalles bioquímicos y biofísicos de dichos procesos. Los resultados y descubrimientos comentados en este artículo permitirán a muy corto plazo una interpretación correcta de la patofisiología de los pigmentos biliares y una mejor comprensión de los mecanismos responsables de su metabolismo hepático.

BIBLIOGRAFIA COMPLEMENTARIA

BILIRUBIN METABOLISM AND HYPERBILIRUBINEMIA. J. Fevery y J. González-Gallego en *Hepatobiliary Diseases*, dirigido por J. Prieto, J. Rodés y D. A. Shafritz. Springer-Verlag, Berlín, págs. 181-199, 1992.

NEW CONCEPTS ON BILIRUBIN CHEMISTRY, TRANSPORT AND METABOLISM: REPORT OF THE SECOND INTERNATIONAL BILIRUBIN WORKSHOP, APRIL 9-11, 1992, TRIESTE, ITALY. C. Tiribelli y D. J. Ostrow en *Hepatology*, vol. 17, págs. 715-736, 1993.

HEPATIC TRANSPORT AND BILE SECRETION: PHYSIOLOGY AND PATHOPHYSIOLOGY. Dirigido por N. Tavoloni y P. D. Berk. Raven Press, Nueva York, 1993.

Ciencia y empresa

Realidad virtual

De lo pintado a lo vivo

Los abogados de la realidad virtual (RV) prometen una técnica que sumerge a los adictos en mundos sintéticos de convincente ilusión. Proclama que ha encontrado amplio eco en rotativos prestigiosos de los que hemos entresacado, al azar, estos titulares: "La realidad virtual crece por días", "La realidad virtual, por fin, real" y "La realidad virtual entra en casa". A tenor de ello, diríase que la realidad virtual ha superado ya sus dificultades técnicas y se encuentra lista para irrumpir en el mercado.

Pero las cosas lucen de otra manera, si hemos de creer en el informe que acaba de publicar el Consejo Nacional de Investigación de los EE.UU. (NRC), integrado por un núcleo de expertos en computación, ingenieros y psicólogos. Afirman en su dictamen que "existe un profundo hiato entre la técnica disponible y la que sería menester para alcanzar el potencial de los sistemas de RV". Henry A. Sowizral, que dirige un proyecto de investigación de realidad virtual para la compañía Boeing Computer Services en Bellevue, lo ratifica. "Los tres graves problemas a los que debe hacer frente la realidad virtual son prestaciones, prestaciones y prestaciones", repite con sorna para designar las limitaciones de que adolece en cuanto a representaciones, ordenadores y programación.

No es tan simple engañar al cerebro humano para que crea que hay algo allí donde no lo hay. Hasta ahora, el trabajo investigador se ha concentrado en engañar a los ojos. Para ello se requiere, por supuesto, alta resolución, gran angular y proyecciones tridimensionales. Otro requisito previo demandado: los aparatos capaces de seguir la dirección de su mirada. Pero los cascos de realidad virtual hoy en uso, que colocan una pantalla mínima de cristal líquido delante de cada ojo son granulosos y caros. Hasta un millón de dólares por unidad les cuesta al ejército ciertos prototipos que ofrecen la resolución de una pantalla típica de ordenador de sobremesa mirada a una distancia de unos veinte centímetros.



Los mundos imaginarios pueden ser la provincia de los viajeros de la realidad virtual. Pero un informe del Consejo Nacional norteamericano de Ciencia llega a la conclusión de que esta técnica todavía no puede satisfacer las esperanzas del público

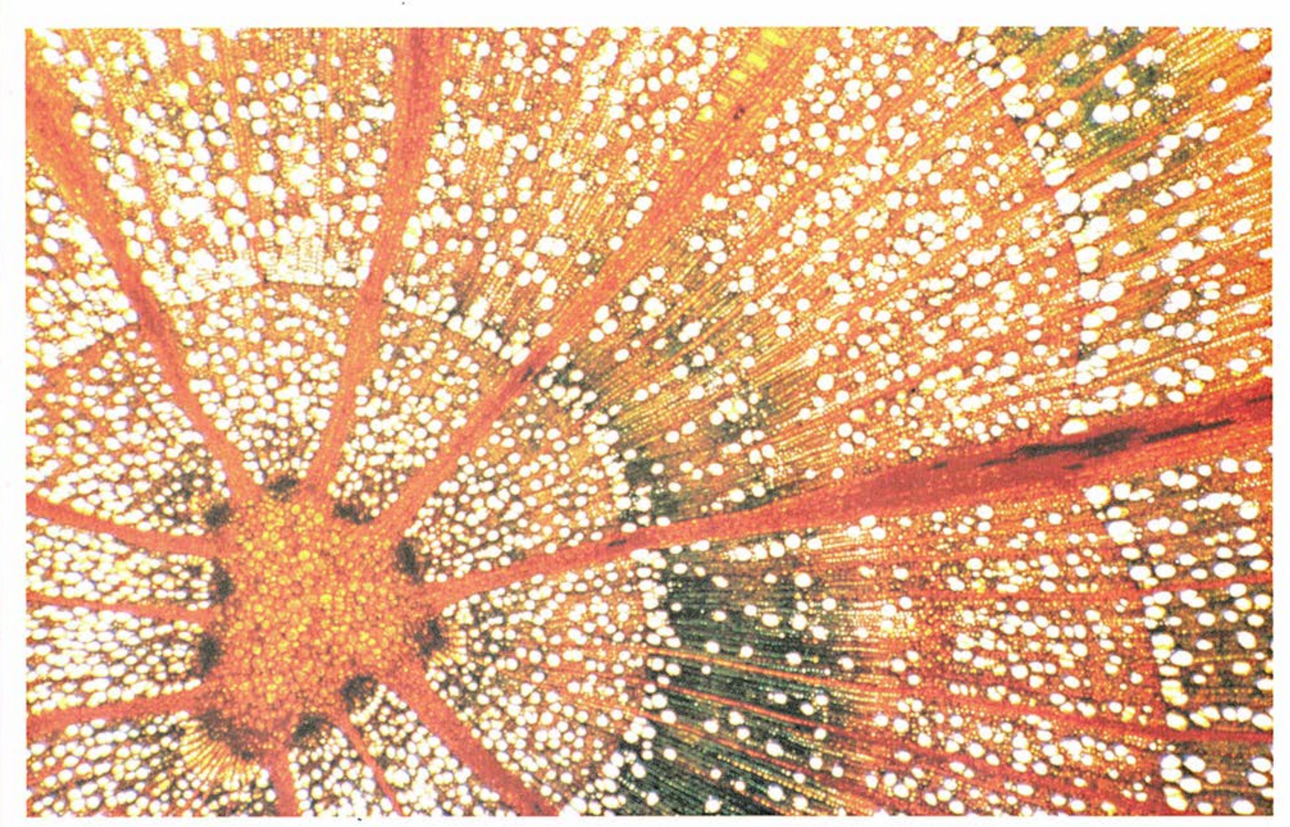
"Los cascos más asequibles vuelven a uno literalmente ciego", dice Sowizral.

Aunque la resolución de las pantallas progresará a buen paso, no será fácil fabricar cascos más ligeros, auténtica pesadilla de los exploradores de mundos virtuales. Cuando el yelmo pesa alrededor del kilo, las pantallas montadas en la cabeza dificultan el giro de ésta. Sumado a una ilusión rígidamente visual de movimiento, el peso induce mareo cinético en los portadores. Las náuseas y los dolores de cabeza son sólo el principio, observa el informe del NRC. "Un problema más serio... es el síndrome de agotamiento, que implica fatiga crónica, falta de iniciativa, somnolencia, letargia, apatía e irritabilidad." Esos síntomas persisten largos períodos, aun cuando hayan sido breves los paseos de diversión por el campo de la virtualidad.

Con todo, el nudo gordiano está en los actuales sistemas de seguimiento de la realidad virtual. "El seguimiento es el hijo bastardo del que nadie quiere hablar", dice Sowizral. Los brazos mecánicos unidos a la cara y las manos, aunque rápidos y precisos, estorban más que ayudan, sobre todo si tenemos tapados los ojos. Los sistemas magnéticos que usan sensores a modo de brújulas, muy populares, "se hallan expuestos a interferencias metálicas, de los ordenadores, por ejemplo. Una vez dejé una lata de refresco cerca de la fuente del campo, y salté 15 metros en el entorno virtual", bromea Sowizral.

Aparte de la visión, los entornos virtuales nunca parecerán reales mientras no podamos alargar la mano y palparlos. Se han ensayado varios dispositivos, controlados por ordenador, para simular la fuerza y la textura. Pero, como advierte Sowizral, "hasta que no se resuelvan algunos problemas, eso puede resultar muy peligroso". Crear la ilusión de la presencia de un cuerpo sólido donde no lo hay requiere fornidos brazos robotizados que sigan la mano y resistan donde conviene. "Si damos un puñetazo sobre la mesa virtual, el aparato necesita múltiples motores de varios caballos para conferirle la sensación de

Concurso internacional Nikkon 1994 sobre microcosmos



El corte del tronco de una joven haya ganó el Concurso internacional Nikkon 1994 sobre microcosmos. El microfotógrafo, Jean Rüegger-Deschenaux, de Zürich, tiñó la muestra con crisoldina y azul astral antes de disparar con 40 magníficos aumentos. El concurso se estableció hace 20 años.

golpe sobre el tablero", observa Sowizral. "Pues bien, varios caballos son suficientes para romperle a uno la muñeca si se ha equivocado en el programa. Por eso los aficionados andan con tiento y ensayan con fuerzas mucho menores." Y así los objetos, aparentemente sólidos, se muestran al tacto blandos y suaves.

Pese a dicho tacto esponjoso, los objetos virtuales seguirán pareciendo irrealmente angulosos y reaccionarán de forma rara al tacto hasta que los ordenadores adquieran una mayor potencia. Para los ordenadores que dibujan mundos virtuales, los objetos tridimensionales están constituidos por polígonos bidimensionales. Los expertos estiman que cada toma de una animación de realidad virtual debe contener unos 80 millones de polígonos para parecer fotorrealista. Para mantener la ilusión de movimiento continuo se requieren, por lo menos, diez imágenes por segundo. (Las películas de dibujos corren a 24 imágenes por segundo; la televisión emplea 30.) Por tanto, cualquier sistema virtual que aspire al realismo visual habrá de calcular y dibujar al menos 800 millones de polígonos por segundo.

Por unos 30 millones de pesetas, la empresa californiana Silicon Graphics ofrece un superordenador gráfico, llamado RealityEngine^2, capaz de producir dos millones de polígonos por segundo en condiciones ideales: 0,4 por ciento de la velocidad necesaria para la verosimilitud. El PixelFlow, sistema experimental más costoso que se halla en fase de construcción por la Universidad de Carolina del Norte en Chapel Hill, se espera alcance unos 30 millones de polígonos por segundo. Virtuality, videojuego anunciado como realidad virtual por W Industries de Leicester, produce escenas de sólo unos cientos de polígonos con toda la complejidad de los mundos fantásticos.

Las aplicaciones útiles de la RV necesitan algo más que imágenes bonitas en movimiento. Los objetos virtuales deben también imitar el comportamiento de sus contrapartidas reales, lo que significa hacer millones de cálculos adicionales cada segundo para garantizar que actúan como sólidos con masa, y no como superficies carentes de la misma. Añádase la sensación al tacto, que muchos programas se esfuerzan por conseguir, y la cantidad de trabajo aumentará de forma espectacular, pues las texturas deben actualizarse cientos de veces por segundo para semejar la realidad viva.

Denuncia el informe emanado del NRC que existe una descompensación entre equipos y desarrollo de la programación; la investigación en equipos de presentación de la realidad virtual crece con rapidez, dejando muy atrás el progreso de los programas. Los estudios cognitivos han puesto de manifiesto el estado de confusión que se crea al separar visión, sonido y tacto de un suceso en unas décimas de milisegundo. Los investigadores en realidad virtual tienen todavía que elaborar los programas operativos capaces de garantizar respuestas simultáneas de las presentaciones visual, auditiva y táctil. Y convertirse en Creador es un papel arriesgado: "Se tardan meses o años en producir esos entornos", dice Sowizral.

¿A qué viene, pues, todo el revuelo levantado con la supuesta llegada de la realidad virtual, cuando el propio Larson-Mogal estima que habrán de transcurrir ocho o diez años antes

de que las limitadas aptitudes de una RealityEngine2 de realidad virtual se nos ofrezcan desde las estanterías de los mercados? Parte de esa idea errónea obedece a la manera de trabajar de los científicos, más proclives a concentrarse en refinamientos del futuro que en mejorar las limitaciones actuales. Culpa tienen también los medios de comunicación, prontos a la hora de tomar las hipótesis como hechos. "La realidad virtual tiene hoy día una relación de palabras a trabajo altísima", se queja el informe del NCR, que se lamenta también de que la mayoría de los investigadores se interesen, de manera casi única, por los programas de representación gráfica.

Lo que es más precupante: el estudio observa que los expertos involucrados parecen haber abandonado la objetividad científica. "No se ha evaluado con seriedad el grado de utilidad práctica de la realidad virtual", concluía el comité. A la hora de calibrar la relación de rendimiento a inversión en un sistema virtual, suele despreciarse la solución conservadora en aras de otra opción de alta tecnología. Hacer que los estudiantes naden con delfines virtuales suena refrescante, pero llevarlos a un acuario de verdad puede ser más barato y eficaz. Algunas veces la realidad no muerde.

W. WAYT GIBBS

Policía de policías

Redes neuronales en las jefaturas

Las películas americanas nos han familiarizado con la policía de Chicago y su lucha contra la mafia. Venía a ser algo así como la quintaesencia de los guardianes de la sociedad. Pero también ha pasado sus altibajos. Y, lo mismo que en cualquier cuerpo social, también en ella maduran manzanas podridas. La división de asuntos internos, que entiende en alegaciones de faltas cometidas por sus propios policías, reclutó la ayuda de un paquete de programas que trata de emular la manera de operar de las células del cerebro. Cada noventa días, la división se propone publicar una lista de los guardias que el programa sugiere pueden meterse en dificultades.

Para saber quiénes, el programa usa un modelo predictivo llamado red neuronal. El programa predice si es probable que algunos de los 12.500 guardias de la plaza se comporten de manera similar a la de los 200 miembros que fueron expulsados o renunciaron (al hallarse bajo investigación) durante los últimos cinco años por acciones que variaban desde la insubordinación hasta la conducta criminal. Los policías nombrados en las listas de "problemáticos" deben someterse a un curso de reeducación.

La manifiesta capacidad del programa de red neuronal para extraer conclusiones con sentido a partir de datos dispares ha dado por resultado su empleo para múltiples aplicaciones, desde la predicción de recaídas en criminales en libertad condicional hasta la identificación de mosquitos por el zumbido de sus alas. La policía de Chicago puede convertirse en pionera del recurso a la tecnología para anticiparse a las fechorías de los guardianes de la ley.

La red consta de unos programas de simulación de una red de procesadores interconectados. Los elementos procesadores y las conexiones entre ellos corresponden aproximadamente a las neuronas y sinapsis del cerebro. A imagen del cerebro, la red debe someterse a un proceso de "adiestramiento". En la red de la policía, los procesadores de entrada aceptan información personal acerca de un determinado policía —quejas de los ciudadanos y accidentes de tráfico, por ejemplo— que se traduce en una serie de valores numéricos. Esas variables alteran la intensidad de las señales, o pesos sinápticos, que pasan de un procesador a otro. La variación de los pesos pone en marcha una cadena de sucesos: la intensidad de las señales se multiplica por otros valores y se suma a ellos en cada procesador. El proceso continúa hasta que la red proporciona valores para estimar la probabilidad o no de expulsión. Los valores se comparan después con otro número, un cero o un uno, que representa si el policía en cuestión ha sido en efecto expulsado o goza de buena reputación.

Si la red ha hecho un pronóstico incorrecto, y generalmente lo hace al principio, una fórmula matemática corrige los pesos. Pasando por la red cientos de ejemplos de policías expulsados y otros con buena hoja de servicios, la red ajusta continuamente los pesos durante una media hora. Con ello "aprende" a hacer predicciones sistemáticamente correctas.

Por lo menos, así es como se supone que funcionan las cosas. Pero las redes neurales tienen sus agnósticos declarados. El primero, el sindicato de policía. Cuando el sindicato se enteró de que había un cerebro computarizado inspeccionando los archivos de personal para descubrir los policías conflictivos, sufrió el equivalente institucional de un aneurisma. "Es absolutamente ridículo; apesta", exclama furioso Bill Nolan, presidente de la Agrupación Solidaria de la Policía de Chicago. Nolan dice que la red neuronal, a la que llama con desprecio "esa bola de cristal", no es más que una táctica del departamento para evitar meter en cintura a sus policías. "¿Tiene Vd. un tío que no cumple? Pues supervíselo, corríjale", pide Nolan. Y añade: "Les dije que si este trasto es tan bueno, tendríamos que dárselo a todos los inspectores, para que resolvieran todos los robos y asesinatos."

Las impresiones de Nolan no difieren, en esencia, de las expuestas por algunos psicólogos del conocimiento y científicos de la computación. "Vudú", sentencia Zenón Pylsyhyn, profesor de ciencia cognoscitiva en la Universidad Rutgers. "La gente está fascinada ante la perspectiva de obtener información por medios misteriosos a lo Frankenstein: ¡por vudú! Y ha habido pocos intentos de hacer esto con tanto éxito como con las redes neuronales."

Objetan los críticos que las redes neuronales son una especie de caja negra: no indican cómo llegan a una conclusión. A diferencia de los sistemas expertos —otra clase de técnica de inteligencia artificial que hace recomendaciones basándose en un conjunto explícito de reglas—, las redes neurales operan mediante complicados procesos no lineales. "La potencia de la red neuronal, como tal, reside en los pesos de conexión, una vasta tabla numérica que desafía un análisis efectivo", escriben Charles X. Ling y A. K. Dewdney, de la Universidad de Ontario Occidental. "Es casi imposible interpretar y comprender lo que las redes neuronales de tamaño moderado han aprendido. En cuanto técnica, el método merece atención, pero en cuanto ciencia eso solo basta ya para denunciar su fracaso."

Ling y Dewdney representan un punto de vista en un debate que ha durado decenios. Dejando aparte la hipérbole acerca de las semejanzas con el cerebro humano, los defensores de las redes neuronales sostienen que la técnica no es otra cosa que un complicado giro del método estadístico normal de deducir un comportamiento a partir de valores numéricos trazando una curva a través de un conjunto de puntos. Además, las metodologías subyacentes están abiertas al análisis. "Durante la ma-

yor parte del tiempo, los usuarios no necesitan saber qué está haciendo el programa; sólo necesitan saber si va bien", aduce Michael Mitman, de California Scientific Software, que vendió al Departamento de Policía de Chicago su paquete informático BrainMaker Professional.

Pero lo cierto es que la división de asuntos internos descubrió que la mitad de los 91 individuos identificados por la red neuronal habían sido ya asignados a un programa de rehabilitación por el departamento de personal, denunciados por su conducta punible. Se trata ahora de que la informática complemente ese programa, dejando a la división de asuntos internos que identifique a los policías proclives a la fechoría antes de que la cometan y pasen a otras instancias judiciales. "En compañías de 150 hombres, no sería necesario aplicar esa técnica ", sostiene Raymond Risley, de la división de asuntos internos. "Pero en el caso de la policía metropolitana de la ciudad de Chicago, resulta casi imposible identificar a todos los individuos en peligro."

GARY STIX

Prehistoria informática

Réplicas tridimensionales

La Edad de Piedra se da literalmente la mano con la era espacial en el laboratorio de John Kappelman, de la Universidad de Texas en Austin. Allí, los haces de láser se reflejan en cráneos y los rayos X penetran a través de huesos antiquísimos. Un ordenador sigue los resultados y compila la información concerniente a la morfología exacta de los especímenes en tres dimensiones. Los restos igual pueden ser de esqueletos de aborígenes norteamericanos muertos hace mucho tiempo que fósiles de homínidos remotísimos. En cuestión de horas, tales especímenes se transforman en archivos de datos, que se almacenan en un CD-ROM o se convierten en réplicas cabales del original. "Es una tecnología novedosa, apenas experimentada", explica Kappelman ilusionado. El antropólogo está convencido de que el proceso transformará su disciplina y resolverá algunos de los conflictos enojosos que han surgido a propósito del derecho de propiedad y del derecho de investigación.

No es frecuente tamaña inmersión en el mar de la alta tecnología para una disciplina acostumbrada a desenvolverse con el cuaderno de campo y el pie de rey. "Los arqueólogos y los antropólogos siempre acceden a los progresos técnicos veinte años después que todos los demás", confiesa Tomás R. Hester, director del Laboratorio de Investigación Arqueológica de Texas. Lo que no se cumplió en Kappelman. Hace pocos años, cayó éste en la cuenta de que la electrónica podía ofrecerle una manera más directa y completa de analizar los fósiles. No tuvo que mirar muy lejos para madurar la idea. "Nos encontramos en el corazón de 'Silicon Hills'", dice refiriéndose a la concentración de empresas de alta tecnología que rodean a la ciudad de Austin.

Presintiendo una buena oportunidad para promocionar sus productos y explorar nuevos mercados, tres compañías —Digibotics, Scientific Measurement Systems y DTM Corporation— se aprestaron a colaborar con el grupo de Kappelman. Cada empresa proporciona una pieza complementaria de equipo. Digibotics fabrica equipos láser de exploración



tridimensional, que registran los contornos de una muestra haciendo incidir un haz láser sobre su superficie. Los equipos están preparados para captar detalles de diámetro inferior a un milímetro.

Scientific Measurement Systems vende máquinas de tomografía computarizada, técnica de construcción de imágenes tridimensionales por rayos X. Los aparatos tomográficos de la última generación proporcionan suficiente resolución para descubrir desde el esmalte gastado en el diente de un homínido hasta las sutilezas de torneado en la alfarería de un cuenco, pasando por las fracturas soldadas en un hueso prehistórico.

Tanto la exploración láser como la tomografía computarizada producen archivos digitales de datos que describen la forma del espécimen. "Pero todavía somos animales táctiles", reflexiona Kappelman. "Millones de años de evolución nos han enseñado a aprender tocando." Para satisfacer esa necesidad, recurrió a DTM Corporation, que ha desarrollado la aglutinación por láser: nuevo proceso que consiste en una especie de fotocopia en tres dimensiones. En esencia, la aglutinación ("sintering") invierte el resultado de la exploración láser: un haz láser guiado por ordenador perfila la forma de la muestra y la esculpe en polvo de nylon o policarbonato. El láser funde después el polvo y fabrica un duplicado.

Desconocidas para la mayoría de los antropólogos, las técnicas de exploración por ordenador y de replicación resultan todavía caras. Stephen Koch, presidente de Digibotics, cifra en cinco millones de pesetas el valor de un equipo de exploración láser, incluido el descuento que se haría para los centros universitarios. Y subiría a 130 millones de pesetas un conjunto completo de exploración, tomografía y aglutinación. De momento, el laboratorio de la Universidad de Texas en Austin ha comprado ya el explorador láser.

Hay que dedicar algún tiempo al proceso de aprendizaje de aplicación de las técnicas de análisis ingenieril a la investigación antropológica. "De aquí a medio año sabremos qué puede dar de sí la nueva técnica", nos comenta Samuel Wilson, que colabora con Kappelman. Más escéptico se muestra Ralph L. Holloway, antropólogo que trabaja en la Universidad de Columbia: "No es algo forzosamente necesario ni apremiante. No queremos acabar atrapados en un sistema que le obliga a uno a invertir bastante tiempo."

Si las nuevas tecnologías cumplen cuanto prometen, podrían ayudar a los museos y universidades a repatriar los restos de los indios aborígenes. Con la aprobación en 1990 de la Ley de Protección y Repatriación de Tumbas de Aborígenes Norteamericanos, todas las instituciones, públicas y privadas, que aceptan dinero federal están obligadas a cumplir las peticiones de devolución y nuevo enterramiento de huesos y pertenencias. En cumplimiento de la norma se ha emprendido un gigantesco programa de catalogación de tales restos que ha de concluir antes del mes de noviembre de este año de 1995.

Pero la ley ha provocado también preocupación entre los antropólogos sobre la enorme cantidad de conocimientos que desaparecerá sepultada bajo el suelo. "Una vez enterrado, ha desaparecido para siempre", dice Kappelman.

La exploración láser podría ofrecer una manera de devolver los objetos a sus propietarios legales conservando al mismo tiempo un simulacro electrónico (o un modelo real) para futuros estudios. "El proceso no es en manera alguna destructivo", insiste Kappelman. "Solamente se trata de iluminar el objeto." Al menos una tribu, que ha solicitado se mantenga el anonimato, está de acuerdo: consintieron en que el grupo de Kappelman explorase los huesos antes de ser reenterrados. La Institución Smithsoniana está también experimentando con la exploración láser. Pero de nuevo surge el problema del dinero, ya que muchos museos han empeñado su presupuesto en la ley de repatriación. Martha Graham, del Museo Americano de Historia Natural, confiesa que al centro "no le queda ni un duro para trabajos distintos de los comprometidos en el cumplimiento de dicha norma".

La repatriación es sólo una aplicación potencial. Con el hábil manejo de la técnica, podrían revolucionarse la enseñanza y la investigación, alterando las reglas sobre las personas con derecho al acceso a los materiales primarios. En la Universidad de Texas, y con ayuda de una subvención de la Fundación Nacional de Ciencia, se está compilando una base de datos en CD-ROM que elaborará imágenes de artefactos raros, fósiles frágiles y primates extintos. Los estudiantes ayudarán en la tediosa tarea de la exploración láser. "Tenemos montones de alumnos dispuestos a trabajar aquí; estaremos explorando casi las 24 horas del día", dice Kappelman.

¿Y después del proyecto en CD-ROM? "Con el tiempo, habrá un archivo Internet. Tardaremos todavía algunos años, pero será una derivación inevitable." Si el precio de los equipos de aglutinación baja, muchas instituciones podrían adquirir los equipos y enlazarlos a un ordenador conectado a la Internet. Un antropólogo podría entonces pedir un archivo a través del módem, cargarlo y después imprimir una perfecta reproducción de un fósil raro.



La exploración láser capta una imagen tridimensional de un hueso pelviano de "Lucy", homínido primitivo. El hueso explorado se puede después analizar, manipular o replicar en el ordenador

COREY S. POWELL

Libre de contaminación

El lago Baikal

Durante décadas, el lago Baikal simbolizó en Rusia la amenaza del desarrollo económico contra la naturaleza. Este reservorio de agua dulce, el mayor, el más antiguo y el más profundo del planeta, ha cautivado el corazón de los poetas y la mente de los científicos, preocupados éstos por su contaminación creciente. Ahora hay razones para esperar que el lago Baikal signifique algo más: el compromiso racional entre las necesidades económicas de un pueblo y las necesidades ecológicas de su tierra.

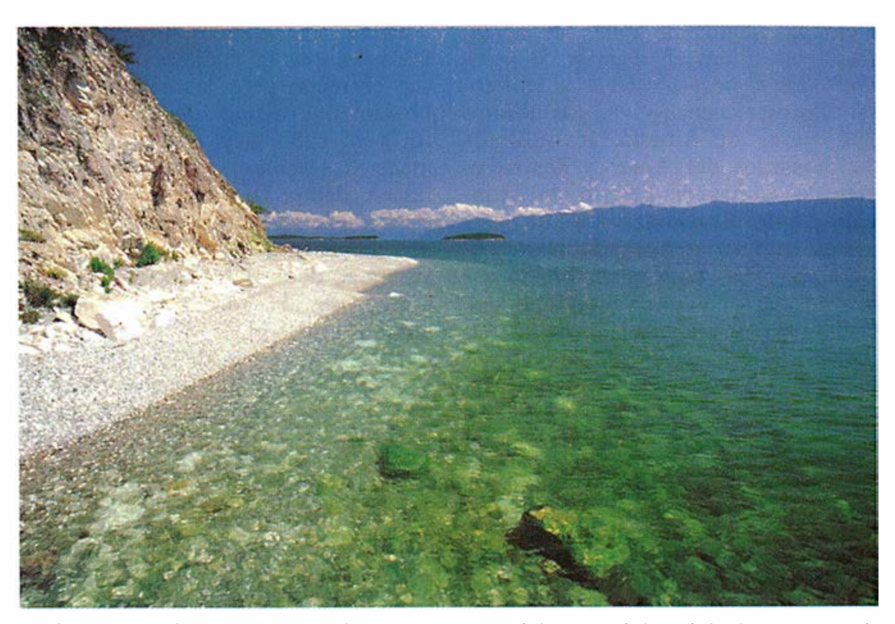
Con unas migajas de apoyo por parte de los Estados Unidos y de las Naciones Unidas, un equipo de ecólogos norteamericanos ha persuadido a Rusia y a Mongolia para que desarrollen programas de uso viable del suelo en la cuenca hidrológica del Baikal. Se proponen salvar el lago y conducir la región hacia una economía de mercado.

Aunque su éxito no está asegurado, los acuerdos del Baikal comienzan a tomarse por modelo en otros lugares. En noviembre del año pasado, China y Rusia empezaron a bosquejar un plan similar para la cuenca del río Ussuri. La república de Altai, en Siberia, también ha aceptado trabajar en una "zona ecológica y económica".

Los proyectos comparten un enfoque y un director comunes: George D. Davis, quien ha proseguido con una estrategia que le reportó éxitos veinte años atrás en la protección del parque neoyorquino de Adirondack, con una superficie de 2,5 millones de kilómetros cuadrados. Inspirado por las leyes de zonificación que las ciudades emplean para separar de las áreas residenciales las industriales, Davis ordenó un estudio para determinar la capacidad de carga de los recursos del parque. Después, redactó las normas y dibujó los mapas de zonificación que indicaban dónde permitir la explotación forestal, la agrícola y ganadera, así como la construcción de viviendas; y ello, con la dificultad añadida de que el 58 % de la reserva era de propiedad privada.

Tras recabar una subvención de la Fundación MacArthur en 1989, Davis fue invitado a aplicar su método a la cuenca hidrológica del Baikal, una región de 60 millones de kilómetros cuadrados que se extiende por Mongolia y Rusia. Con esa financiación, Davis se rodeó de un equipo de 30 expertos, norteamericanos y rusos.

Dividieron arbitrariamente la cuenca hidrográfica en 25 tipos de zonas distintas, que iban desde el suelo agrícola hasta los parques industriales. A cada zona le asignaron usos "preferentes" y "condicionados"; estos últimos requieren permiso. Todo lo que no esté especificado está prohibido. Más de 21 millones de kilómetros cuadrados, incluido el propio lago, se han segmentado en parques nacionales, reservas científicas, paisajes, ríos pintorescos, cinturones verdes y lugares de interés.



El lago Baikal, que contiene la quinta parte del agua dulce del planeta, puede ser salvado por la zonificación

Para el Baikal, la protección llega en el momento justo. Con más de 1600 metros de profundidad y rebosante de oxígeno, constituye el hogar de unas 1800 especies endémicas. Sus 25.000 kilómetros cúbicos de agua potable son tan puros como el rocío. Pero la lluvia que ahora cae es ácida, contaminada por las chimeneas de Irkutsk, al oeste. Muchos más contaminantes le llegan desde el río Selenka.

Gary A. Cook señala que el riesgo ha remitido notablemente con la paralización de la industria rusa. Aunque no está claro si los nuevos propietarios de fábricas y explotaciones agrarias, antaño estatales, cumplirán los planes para el Baikal. En 1993, después que la república de Buriato y Chita Oblast adoptaran la estrategia de zonificación, el presidente ruso Boris Yeltsin firmó un decreto por el que se creaba (pero no se financiaba) una comisión para llevarla a cabo.

Hasta ahora, dice Sergei G. Shapjaev, "no hemos encontrado ninguna oposición organizada. El problema más grave parece ser la falta de mecanismos reales que obliguen a cumplir la ley." Davis confía en que, puesto que los siberianos obtuvieron el derecho de poseer tierras en fecha muy reciente, "no van a notar el aguijón de las restricciones sobre lo que pueden hacer con ellas". Más difícil de apaciguar parece ser el comercio. La Oficina norteamericana para el Desarrollo Internacional ha prometido 3,4 millones de dólares para 12 proyectos dedicados al ecoturismo y la gestión forestal.

Las dos provincias que han adoptado el programa de uso del suelo cubren el 95 % de la cuenca hidrográfica rusa. Pero Irkutsk, que se ha resistido al plan del Baikal, supone el 40 % de la línea de costa del lago y gran parte de los desechos que a él se vierten. Otros 28 millones de kilómetros cuadrados de la cuenca hidrográfica se encuentran en Mongolia.

Tanto o más importante que el proyecto del Baikal es la celeridad con que se le está copiando. Davis trabaja ahora con los territorios de Jabarovsk y Primorsky, a 2400 kilómetros al este del Baikal. Encajados en la provincia china de Heilongjiang, con la que comparten el río Ussuri, estos rusos del Lejano Este están menos preocupados por la amenaza de la contaminación que por la tentación de vender los derechos de sus frondosos bosques, donde halla cobijo la mayor diversidad de especies vegetales de toda la antigua Unión Soviética.

"En el lado chino, la cuenca del Ussuri contiene la mayoría de los humedales que quedan en todo el país", dice Jim Harris. "Se han dre-

nado ya cientos de miles de hectáreas, convertidas en campos de cultivo." Vale la pena preservar lo existente. "En esta cuenca viven los últimos 250 tigres de Siberia, los últimos 30 leopardos del Amur y dos especies de grullas en peligro de extinción", apostilla Davis.

Los equipos científicos ruso, americano y chino presentarán pronto sus recomendaciones para asegurar que el desarrollo de una región de 24 millones de kilómetros cuadrados no sobrecargue a los ecosistemas. Davis señala que los planes han generado poca oposición debido a que se trata de un área muy remota y al lento crecimiento de su economía. "Tenemos la gran suerte de que en ninguna de estas áreas hemos de enfrentarnos a una situación de superpoblación. Y si podemos demostrar que puede funcionar en estas regiones, habrá que considerar entonces otras zonas más complejas." Parece como si a Davis se le ofreciera esta oportunidad: ha sido requerido su asesoramiento por los indios haisla, de la Columbia Británica, los indios misquitos de Nicaragua y funcionarios de Bolivia y Chile.

W. WAYT GIBBS

El platillo de la balanza

Polémica entre matemáticos

A los matemáticos les gusta creer que sus verdades son objetivas. Puede que sea así, pero una disputa, persistente y llena de rencor, acerca de un famoso problema, la conjetura de Kepler, ha descubierto cuán subjetivo puede ser el proceso de juzgar esas verdades. La polémica empezó de una manera inocente, hace cuatro años, cuando Wu-Yi Hsiang decidió dar un curso sobre geometría clásica. Para afinar sus habilidades en esta disciplina, tomó una conjetura formulada en 1611 por Johannes Kepler, el mismo polígrafo alemán que descubrió que las órbitas de los planetas eran elípticas.

Kepler propuso que el método más compacto de amontonar esferas es el que la naturaleza aprovecha para disponer los átomos en los cristales y los fruteros para apilar naranjas en pirámides de cuatro caras. La manera más sencilla de crear ese patrón es formar una capa de esferas hecha de filas horizontales y verticales regulares y colocar las esferas de la capa siguiente en los intervalos que quedan entre cada cuatro esferas de la capa inferior.

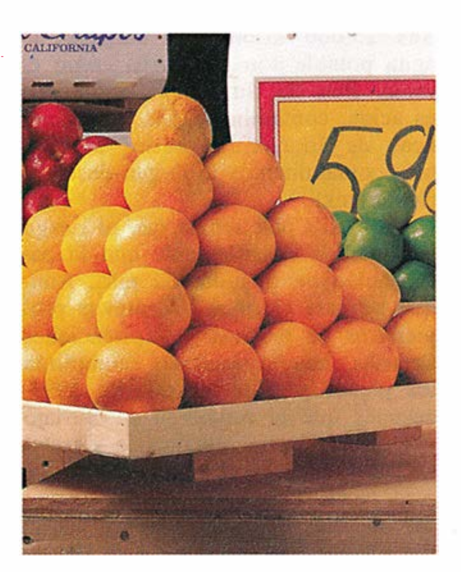
Pocos matemáticos tienen alguna duda de que la conjetura de Kepler sea correcta, pero les resulta endiabladamente difícil demostrarla. Al fin y al cabo, hay una infinidad de maneras de colocar esferas en un volumen dado. Douglas J. Muder ha establecido que las esferas no pueden llenar más que un 77,3 % del volumen, pero la conjetura de Kepler afirma que la cota superior es de alrededor del 74 %.

Tras seis meses de meditar sobre el problema, Hsiang se convenció de que tenía una prueba. Aunque su argumento se basaba en técnicas hasta cierto punto corrientes de la geometría y del cálculo, era larga —el primer borrador ocupaba más de cien páginas— e intrincada. El enfoque de Hsiang consistía en calcular la densidad "local" de varias configuraciones finitas de esferas y extrapolar esos resultados a volúmenes infinitos.

Hsiang puso en circulación un borrador de su prueba. No tardaron en ponerla en entredicho cuatro expertos en empaquetamiento de esferas —Muder, John H. Conway, Neil J. Sloane y Thomas C. Hales. El grupo deploraba que el texto, siendo tan largo, fuese pobre en detalles: los saltos de los casos particulares a las generalizaciones no estaban tan justificados como debieran.

Lejos de retractarse, Hsiang remitió su artículo al *International Journal of Mathematics*. Una vez Hsiang hubo hecho unas revisiones, la revista publicó las 92 páginas del artículo en octubre de 1993.

La primavera pasada, Conway, Hales, Muder y Sloane anunciaron en *The Mathematical Intelligencer* que "no consideran que el trabajo de Hsiang sea una prueba de la conjetura de Kepler o que pueda completarse para que llegue a serlo en un tiempo razonable." En el número de verano del *Intelligencer*, Hales presentó un acre resumen de doce páginas de las objeciones que el grupo le hacía a la obra de Hsiang. Sostenía que el artículo no pasaba de ser, en el mejor de los casos, una serie de conjeturas que, si se demostrasen, podría ser una demostración.



Estas pirámides de fruta exhiben la que, según conjeturó Johannes Kepler en 1611, es la disposición de esferas más compacta: la red cúbica centrada en las caras

Conway predice que Hsiang sacrificará su "distinguida" reputación si insiste en decir que tiene una prueba. Según Muder, la polémica ya ha desanimado a otros matemáticos a seguir trabajando en la conjetura de Kepler; nadie quiere dedicarse a un problema resuelto. Sloane contrapone el comportamiento de Hsiang al de Andrew J. Wiles. En 1993, Wiles anunció que había demostrado el "último teorema de Fermat", pero anduvo presto en retirar su pretensión en cuanto los colegas le señalaron ciertas deficiencias.

Hsiang, cuya réplica a los críticos publicará el *Intelligencer* el próximo invierno, responde que las quejas de éstos consisten en "malos entendidos, malas interpretaciones y acusaciones falsas". Su prueba "ofrece todo lo que es imprescindible conocer" y sólo omite "cálculos aburridos". Lo único que Hsiang admite es que puede que haya un "problema de comunicación". En colaboración con el matemático húngaro Karoly Bezdek, planea elaborar una versión más detallada de su prueba.

Bezdek coincide con Hsiang en que Hales y sus colegas "o han entendido mal o deliberadamente no han querido seguir" sus ideas. Sin embargo, cree también que la demostración de Hsiang no es completa. "Soy optimista" en cuanto a que al menos un elemento crucial de la prueba podrá completarse, señala, "pero hay lagunas". Además, reconoce que si se inclina en favor del trabajo de Hsiang, a lo mejor es, en parte, porque extiende un enfoque que se desarrolló originalmente en Hungría.

La verdad es que algunos observadores temen que la pelea en torno a la conjetura de Kepler no sea sino un indicio de que las matemáticas se enfrentan a un dilema más profundo y de mayor alcance: a medida que la disciplina se hace cada vez más compleja y especializada, la evaluación de las pruebas es más y más difícil.

JOHN HORGAN

Juegos matemáticos

Ian Stewart

Ajedrez en un tablero de go

El día había sido largo y lluvioso, y a los trillizos Vascobiaga —Begoña, Maite y Guillermo, alias "Willi"— se les estaban agotando las ideas a gran velocidad.

"Me aburro", suspiró Maite.

"¿Por qué no echamos otra partida de go?", propuso Willi.

"Hemos jugado tanto a eso, que todo cuanto miro se convierte en redondeles blancos y negros", replicó Maite.

"Pues entonces, juguemos al ajedrez", ofreció Begoña. "El Círculo de Juegos y Rompecabezas de la Sociedad Arquimediana —una sociedad de estudiantes de la Universidad de Cambridge, en Inglaterra— ha inventado una forma de jugar al ajedrez en un tablero de go. Bueno, algo parecido al ajedrez", matizó la chica. "No es exactamente lo mismo, tenéis que comprenderlo. La verdad es que si me apuráis mucho diría que es bastante diferente, pero..."

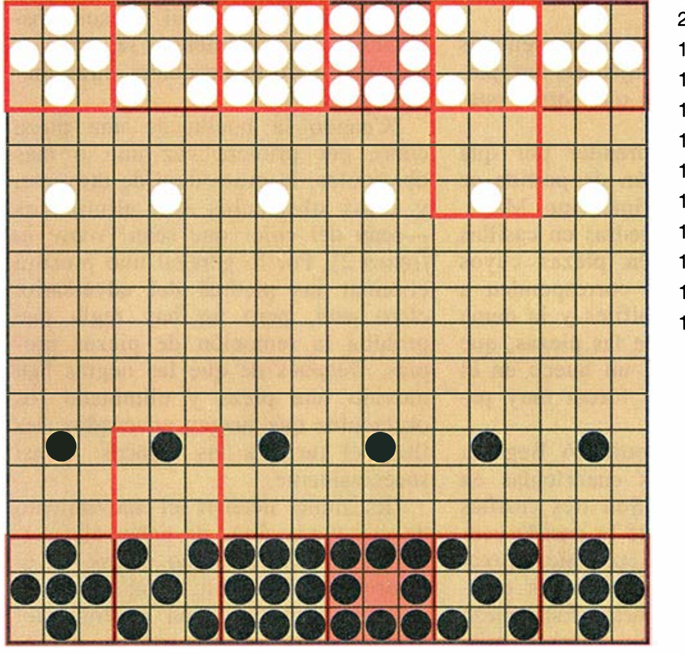
"Vamos, déjate de rodeos. ¿Es divertido?"

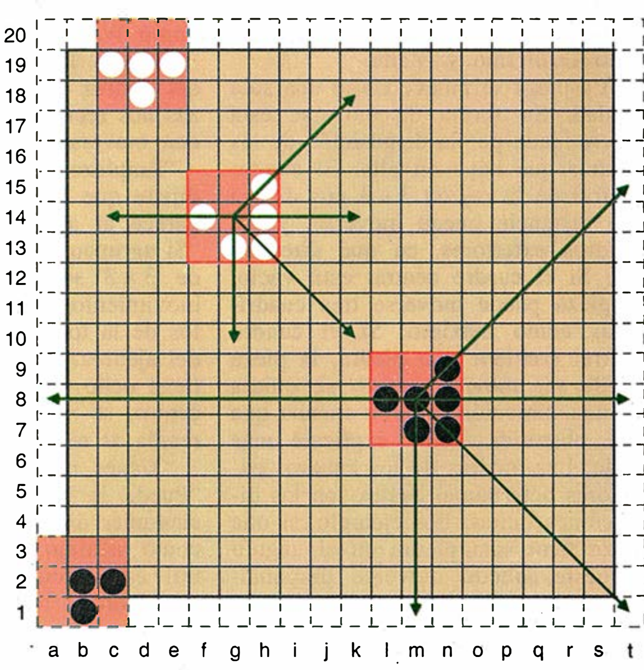
"Sin la menor duda. Es un juego muy poco corriente, como veréis si me dejáis que os lo explique. Durante el juego, las piezas se descomponen, por así decirlo, y vuelven a ensamblarse. El juego se llama 'godrez' porque es un híbrido entre el go y el ajedrez. Se juega al godrez en un tablero de 18 por 18 casillas, que es, casualmente, como están dispuestas las hileras de un tablero ordinario de go. En el go tradicional se ocupan las intersecciones de las líneas del cuadriculado, pero en el godrez se juega en los cuadrados. Para formar las piezas se utilizan las piedras de go, aunque de una forma muy curiosa.

"Hay dos bandos, el de las blancas y el de las negras, cada uno de los cuales dispone de 43 piedras de su color. Juegan alternativamente, por turnos; las negras abren la partida. La posición inicial es casi igual que la de ajedrez [*véase la figura 1, izquierda*].

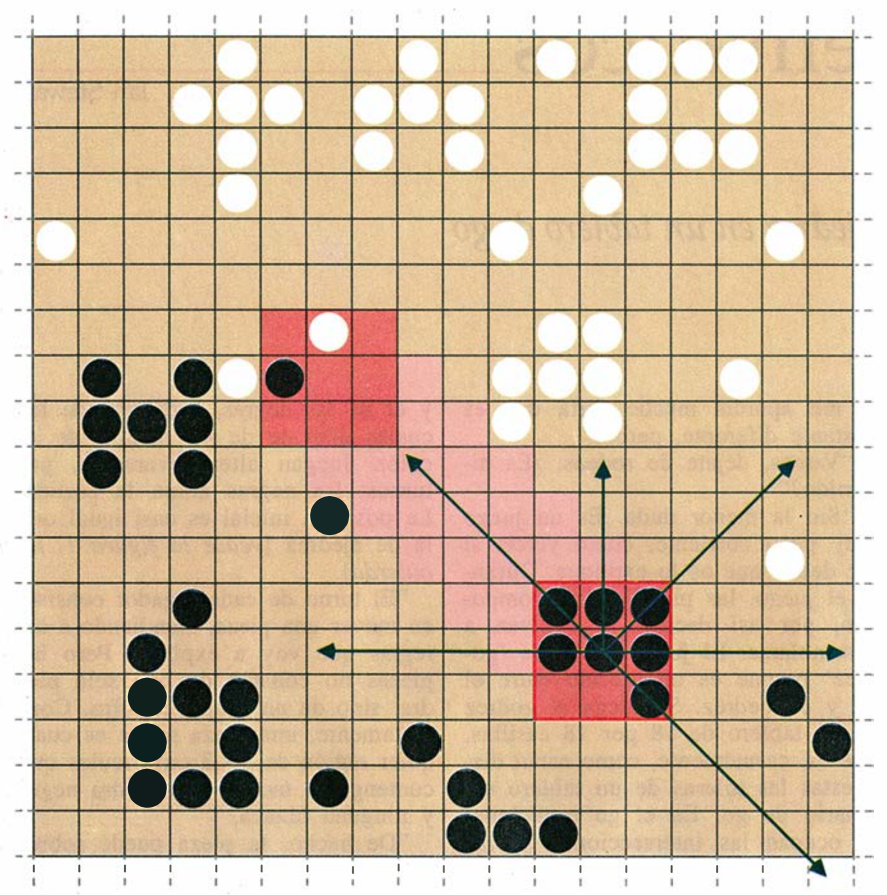
"El turno de cada jugador consiste en mover una pieza, atendiendo a las reglas que voy a explicar. Pero las piezas no constan de una sola piedra, sino de un grupo de ellas. Concretamente, una pieza negra es cualquier región de 3 × 3 cuadrículas que contenga al menos una piedra negra y ninguna blanca.

"De hecho, la pieza puede sobresalir parcialmente del borde del tablero, con tal de que no esté del todo fuera de él [*véase figura 1, derecha*]. Análogamente, una pieza blanca es cualquier región de 3 × 3 que contenga al menos una piedra





*1. Posición de salida del godrez (*izquierda*). El perfil en rojo delinea grupos de piedras cuyos movimientos se asemejan a los de la torre, el alfil, la dama, el rey y el peón. Los aros de los jugadores están sombreados. Una pieza es una región cualquiera de 3 × 3 que contenga al menos una piedra y no las contenga de distinto color. Como se refleja a la derecha, las piedras de una pieza determinan las direcciones en que se puede mover y también la distancia máxima: tres cuadrículas si la casilla central está vacía; una distancia ilimitada si está ocupada*



*2. La pieza tiene que dejar de moverse cuando su huella (*que vemos sombreada arriba*) encuentra por primera vez otras piedras, pero puede detenerse antes. Las piedras cubiertas por su emplazamiento final son retiradas del tablero, sean del color que sean*

blanca y ninguna piedra negra. ¿Me seguís hasta aquí?"

"Pues claro", respondieron al unísono Guillermo y Maite.

"La pieza se mueve como una sola unidad. Su forma de moverse está determinada por la disposición de las piedras que haya en ella. El cuadro central de la región 3 × 3 nos dice a qué distancia puede moverse, y los cuadros exteriores, en qué direcciones. Si el cuadro central está vacío, la pieza puede moverse tres cuadrículas como máximo. Si el cuadro central contiene una piedra, la pieza puede ser movida cuanto se quiera en una dirección legal, a menos que esté obstruida, como explicaré más tarde. Las piedras de los cuadros exteriores determinan cuáles son las direcciones lícitas. Por ejemplo, si una pieza tiene una piedra en el ángulo noroeste, puede moverse diagonalmente en dirección noroccidental. Pero no puede retornar en sentido contrario a menos que tenga una piedra en su ángulo suroriental. Si tiene una piedra en la posición norte, puede moverse hacia el norte, y así sucesivamente."

"¿Por tanto, la forma de la pieza viene a ser como un mapa de lo lejos que puede ir y adónde?", preguntó Willi.

"Así es. Esa es una de las ventajas del godrez sobre el ajedrez: las piezas nos recuerdan de qué forma pueden moverse."

"Empiezo a comprender por qué dijiste que la posición de partida se parece al ajedrez", intervino Maite. "Si agrupamos las piedras en casillas de 3 × 3 se obtienen piezas cuyos movimientos legales corresponden a los de la torre, los alfiles y la dama del ajedrez. Y una de las piezas, que tiene ocho piedras y un hueco en el centro, se mueve de forma muy parecida al rey."

"Tienes razón", confirmó Begoña. "Puede ir hasta tres cuadrículas en cualquier dirección. Son tres casillas como máximo, porque la casilla central está vacía; en cualquier dirección porque todas las casillas exteriores contienen piedras. Esta pieza —o cualquier otra que tenga la misma forma— se llama 'aro', y es importante defenderla."

"¿Por qué?"

"Porque si tu aro resulta dañado, pierdes la partida. Aunque la cosa no es tan sencilla. Durante el juego puedes formar aros extra y, si lo consigues, tu oponente, para vencer, ha de dañarlos todos."

"Imagino que esas piedras individuales situadas delante son como peones", aventuró Guillermo.

"Exactamente. La única pieza del ajedrez que no tiene equivalencia aquí, ni siquiera aproximada, es el caballo. Ello se debe a que en el godrez todos los movimientos son por desplazamiento lineal, nunca a saltos. Ahora, como ya he dicho, supongamos que sea el turno de jugar las negras. Elige una cualquiera de sus piezas, o sea, una región adecuada de 3 × 3. Fijémonos en que las piezas pueden superponerse unas a otras. Eso nada importa; sencillamente, da más variantes al juego. Una vez elegida una pieza, las negras deciden en qué direcciones puede moverse y hasta qué distancia. El negro mueve entonces la pieza en esa dirección, a la distancia que quiera, teniendo en cuenta las limitaciones que esa pieza tenga. Ahora bien, si por el camino hay un obstáculo, las cosas son más complicadas."

"¿Qué quiere decir 'obstáculo'?"

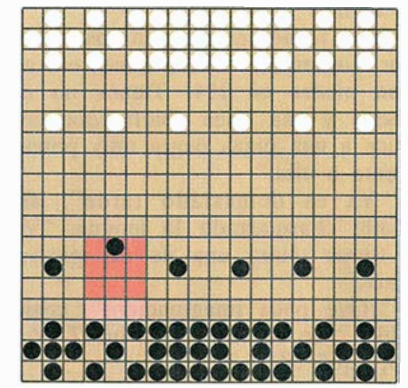
"En primer lugar, definiremos la huella de una pieza como la región 3 × 3 completa ocupada por ella. Al mover, imaginemos que su huella se desplaza conjuntamente con ella. La pieza puede seguir avanzando en su dirección elegida sólo si su huella, al moverse, no tropieza con otras piedras, sean blancas o negras. Las piedras de la pieza móvil no son obstáculos —sencillamente, se mueven juntas—, pero cualesquiera otras piedras sí lo son."

"Cuando la huella de una pieza cubre por primera vez uno o más obstáculos, la pieza deja de moverse, y estos obstáculos son eliminados —sean del color que sean [*véase la figura 2*]. Por lo general uno procura eliminar las piedras del adversario, claro está, pero no hay regla que prohíba la remoción de piezas propias. Después de que las negras han movido una pieza y eliminado los obstáculos que juzgan adecuados, les llega el turno a las blancas. Y así sucesivamente.

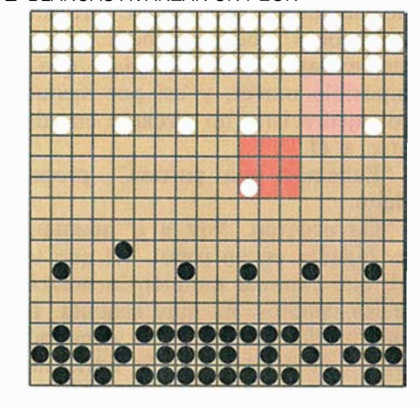
"Es lícito detener el movimiento de una pieza antes de haber alcanzado un obstáculo; como ocurre en el ajedrez, no hay obligación de tomar una pieza al hacer un movimiento. Y una piedra aislada, sita en el centro de un cuadrado de 3 × 3, puede moverse tanto como quiera, pero en ninguna dirección."

"¿Quieres decir que no puede moverse en absoluto?"

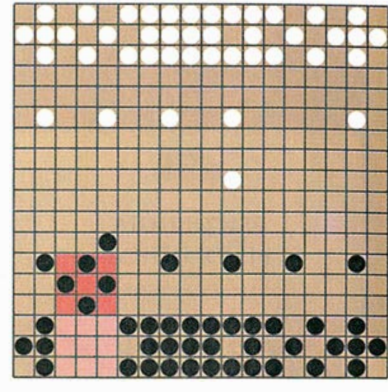
1 ABREN LAS NEGRAS



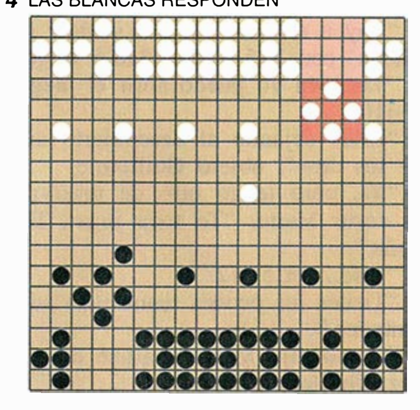
2 BLANCAS AVANZAN UN PEON



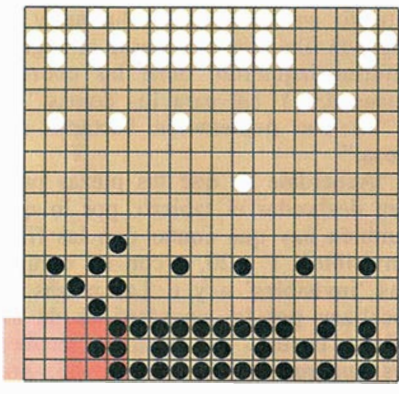
3 LAS NEGRAS SE DISPONE A FORMAR UN SEGUNDO ARO



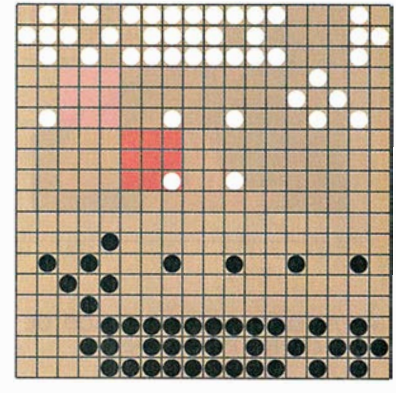
4 LAS BLANCAS RESPONDEN



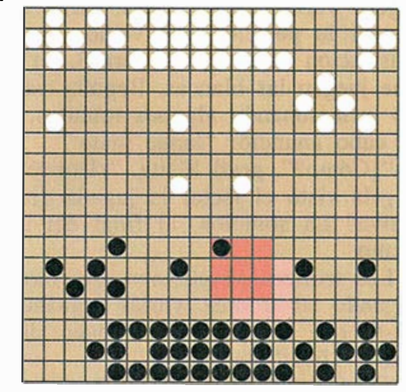
5 LAS NEGRAS FORMAN UN SEGUNDO ARO



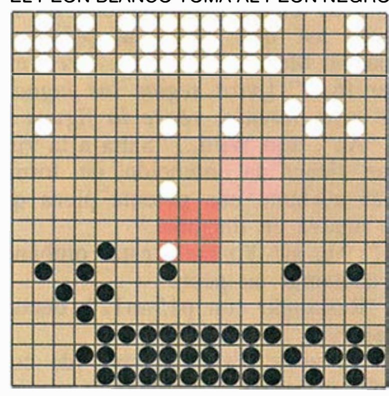
6 LAS BLANCAS AVANZAN AL CENTRO



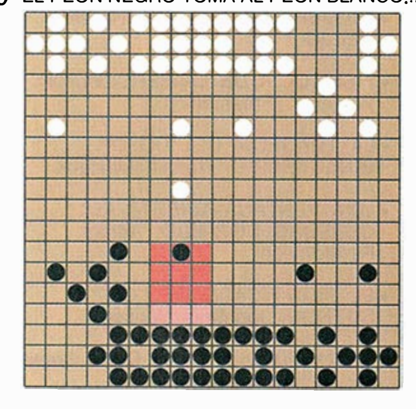
7 LAS NEGRAS SE DIRIGEN AL CENTRO



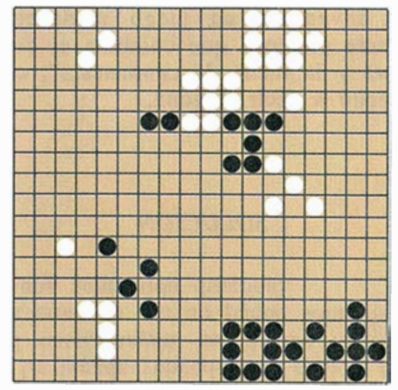
8 EL PEON BLANCO TOMA AL PEON NEGRO



9 EL PEON NEGRO TOMA AL PEON BLANCO...



... MATE EN UNA



3. Las jugadas de apertura de una partida típica de godrez se muestran en los nueve primeros tableros. El sombreado claro muestra las posiciones originales de las piezas, y el sombreado oscuro, su nueva posición. El juego termina en el último de los tableros mostrados: las blancas juegan y dan mate en una

"Lo has entendido muy bien. ¡Ah! Y no están permitidas jugadas que dejen invariable el aspecto del tablero. Como ya he dicho, el juego termina si uno de los jugadores se queda sin aros completos; en ese caso pierde. Por tanto, hemos de proteger al menos uno de los aros propios, exactamente lo mismo que los jugadores de ajedrez han de defender su rey. Pero si durante el juego uno crea aros extra, basta defender uno de ellos. Evidentemente, hay que tener cuidado de no romper un aro propio por desplazamiento de una pieza que se superponga parcialmente al mismo. A menos, repito, que se tenga un aro sobrante completo en algún otro sitio."

"Ya veo," terció Willi. "Aunque la posición de apertura está inspirada en el ajedrez, no es necesario seguir eligiendo las regiones 3×3 de tipo 'ajedrecístico' que determinan la posición inicial. Al mover cuadrados de 3×3 que se traslapen se pueden fragmentar los grupos de piedras originales, y como es obvio, las piezas pueden hacerse con piedras nuevas moviéndose junto a ellas, y pueden también perderlas al ser tomadas por el adversario, o por el bando propio. Por esa razón, el juego resultará bastante diferente del ajedrez."

"Desde luego, a eso se debe que sea entretenido. No obstante, obedece a muchos de los mismos princi-

pios generales: proteger el aro propio, mantener una sólida formación de 'peones', ocupar líneas despejadas, evitar piezas 'clavadas' que es imposible mover, y así por menudo. Los ajedrecistas no tardarán en adaptarse a él."

"Echemos una partida", propuso Maite.

"De acuerdo. Voy a ir apuntando las jugadas en un papel, así que tendré que explicar también la notación de los cuadros. Es muy sencilla. Los cuadros del tablero están numerados verticalmente de 2 a 19; horizontalmente son nombrados por letras, de la b a la s. Por ejemplo, el cuadro situado en la tercera columna y la segunda fila, contando desde el ángulo inferior izquierdo, tiene las coordenadas d3. Y así para los demás."

"Es una numeración bastante curiosa."

"Acuérdate de que las piezas pueden desbordar el tablero, por lo que también hay que contar con las filas 'invisibles' 1 y 20, y con las columnas 'invisibles' a y t. Las piezas se localizan por las coordenadas de su cuadro central. La pieza g5 recubre los cuadros f4, f5, f6, g4, g5, g6, h4, h5, h6."

"Muy bien. Yo jugaré con las negras, así que salgo la primera", anunció Maite. "Vale, pues. Voy a mover f6-f7" [*para seguir las jugadas de esta partida, véase la figura 3*].

"Ah", dijo Willi... "Salta a la vista que tienes la intención de formar más adelante una potente pieza que se mueva en diagonal."

"¿Ah, sí? Yo creía estar probando a adelantar un peón."

"Respondo con p15-m12", declaró Begoña.

"Ya veo... Estás tratando de controlar el centro. Y también te preparas para crear un segundo aro. Pues a eso pueden jugar dos. Muevo e3-e6."

"En efecto, también podemos jugar dos. Yo haré p18-p15."

"¿Y ahora qué hago?", dijo Maite, que no encontraba una buena jugada.

"Prueba b3-e3", propuso Willi.

"¿Está permitido hacer eso? ¡Ah, sí! Las piezas pueden desbordar el tablero... parcialmente. Y así yo formo un segundo aro y refuerzo mi posición."

"Sólo por el momento", replicó Begoña. "Mi jugada es e15-h12."

"Sigues reforzando tu centro, según veo."

"Y a la vez abre una línea para atacar tu aro nuevo", observó Willi.

"Es verdad. ¿Qué te parece m6-l7? ¿Abro así una línea para mí?"

"No está mal."

"Hum... Te voy a apañar yo a ti: m12-j9. Peón por peón. O algo por el estilo."

"Esa jugada no es gran cosa", hizo ver Guillermo. "Ahora tienes clavada una piedra en l14. Tienes un aro detrás y una línea abierta que conduce a una pieza negra muy potente frente a ella."

"Ya es demasiado tarde", dijo Maite chasqueando la lengua. "Voy a tomar ahora una de tus piedras: i6-i7."

"¡Qué poca imaginación!", criticó Willi. "Yo recomendaría h6-i7. Y si no, ¿qué tal j6-k7?"

"Venga, cierra la caja de los piños."

"Voy a reforzar mi aro y abrir mi 'dama' con h15-k12", añadió Bego.

"También yo puedo abrir el juego: i7-i10."

"Estás dejando al descubierto uno de tus aros", objetó Guillermo.

"Ya, pero tengo otro de reserva", le cortó Maite. "Te toca jugar, Begoña."

"Esto... m15-j12."

"Tienes los peones demasiado adelantados."

"Como dijo Maite, ¡cállate de una vez, Guillermo!"

La partida —y la discusión— prosiguió varias jugadas más, hasta alcanzar la posición que vemos en la ilustración final de la figura 3, en la que juegan las blancas. Begoña estuvo reflexionando algunos minutos, y al cabo anunció: "Jaque mate."

"¿Cómo?", gritó Maite. "¿Dónde?"

¿Ha visto el lector la jugada definitiva? La solución, al final del artículo.

Los trillizos Vascobiaga estuvieron jugando al godrez varias semanas, hasta que de pronto, un buen día, Maite dijo: "Me aburro."

"El Círculo de Juegos y Rompecabezas de la Sociedad Arquimediana ha inventado un juego verdaderamente precioso que combina el croquet y la lucha de caimanes...", comenzó a decir Begoña.

SOLUCION

La jugada vencedora es o11-j6. En tal caso, hagan las negras lo que hicieron, su aro no puede sobrevivir a una acometida diagonal.

BIBLIOGRAFIA COMPLEMENTARIA

GESS THE GAME. Paul Bolchover en *Eureka*, vol. 53, páginas 24-27; 1994.

Libros

J. J. Etayo, L. Alonso, E. Farré, A. García-Bellido y R. Margalef

Geometría algebraica

Clásico

METHODS OF ALGEBRAIC GEOMETRY, por W. V. D. Hodge y D. Pedoe. Cambridge Mathematical Library; Cambridge, 1994.

Entre 1947 y 1954 había publicado Cambridge University Press esta obra, que consta de cinco libros distribuidos en tres volúmenes. Posteriormente hizo algunas reproducciones en rústica de la misma y, ahora, vuelve a ponerla en circulación en una edición facsímil, en la seguridad de que los métodos expuestos tienen hoy, todavía, relevancia suficiente.

Apareció en un momento bien perfilado en la evolución de los métodos algebraicos y su aplicación a la geometría. Hay unos, que se consideran ya clásicos, elaborados por la escuela de Enriques, Severi, Castelnuovo, Segre, Bertini, ...y está por entonces al caer la era Grothendieck. Entre ambas etapas se sitúa ese tiempo de extraordinario florecimiento que desde los Noether y van der Waerden a Zariski y Weil conforma toda una época de la geometría algebraica.

Para los lectores que, curtidos en los métodos clásicos de la geometría algebraica, desean adquirir los nuevos y poderosos instrumentos aportados por el álgebra, está pensada la obra. Esto se pone de manifiesto a la hora de elegir el campo base del espacio proyectivo en cuyo seno se sitúan las variedades algebraicas objeto del estudio. Parece que los autores dudaron en tomar, frente al cuerpo complejo sobre el que está construida la geometría clásica, otros cuerpos más generales. Se decantaron por los de característica cero, que, como se sabe, siguen el mismo patrón que el de los complejos, lo que puede facilitar el paso de unos métodos a otros sin el salto que supondría abordar las geometrías sobre cuerpos de característica positiva.

El primer volumen está destinado a proveer de los materiales algebraicos y geométricos que se usarán a lo largo de todo el trabajo. Consta de dos libros, el primero dedicado a los preliminares algebraicos, linealidad y dependencia algebraica, con la novedad de desarrollar esta última a través de la teoría algebraica de los jacobianos. El segundo libro comprende el estudio de las propiedades básicas del espacio proyectivo *n*-dimensional, y lo hace desde los dos puntos de vista, el algebraico y el sintético. Particular interés despierta en cualquier geómetra clásico, el largo capítulo dedicado a la geometría proyectiva sintética y al modo de trasladar la geometría al álgebra y viceversa. O la versión que en uno y otro lenguaje se da de una determinada propiedad: el lema de Pappus, por ejemplo, que es equivalente en el terreno algebraico a que el cuerpo base sea conmutativo, en el geométrico supone que una proyectividad entre rectas está determinada por dos ternas de puntos correspondientes. Termina con el estudio de colineaciones y correlaciones.

También el segundo volumen consta de dos libros. El libro III trata de los principales métodos usados en el desarrollo de la teoría de las variedades algebraicas en un espacio proyectivo *n*-dimensional. Encontramos en él las nociones de irreducibilidad, punto general, dimensión, orden, espacio tangente, etc.; las de correspondencia algebraica y multiplicidad, recogiendo ideas de la geometría italiana, lo mismo que en el capítulo de las intersecciones de variedades algebraicas y sistemas de variedades se evocan los primeros estudios de sistemas de curvas sobre una superficie algebraica que introdujo aquella escuela. Las ideas de van der Waerden, Zariski y Weil, y también de Chow y Chevalley, inspiran el contenido de todo este libro III.

El IV, dedicado a algunas variedades concretas, pretende ilustrar sobre qué propiedades de las mismas pueden ser establecidas a partir de los métodos desarrollados en el libro anterior. Las variedades estudiadas son las cuádricas y las grassmannianas; y en particular las de Schubert, como subvariedades de estas últimas, con su incidencia en las aplicaciones a la geometría numerativa.

El libro V, finalmente, ocupa todo el tercer volumen y está dedicado a la geometría birracional. Tanto los preliminares, teoría de anillos e ideales, como sobre todo el capítulo más central, la teoría de valoraciones, están desarrollados con vistas a las aplicaciones geométricas estrictamente y no a otros campos, como podría ser la teoría de números. Está basado en la obra de Krull, como lo está en la de Zariski el último capítulo sobre transformaciones birracionales. Hay una exposición del problema de resolución de singularidades en superficies algebraicas y en variedades tridimensionales pero faltan aún ocho años para que Hironaka presente, con nuevas técnicas, soluciones al problema general. (J. J. E.)

Del ADN recombinante

A la reacción en cadena de la polimerasa

GENETICS AND MOLECULAR BIOLOGY, por Robert Schleif. The Johns Hopkins University Press; Baltimore-Londres, 1993.

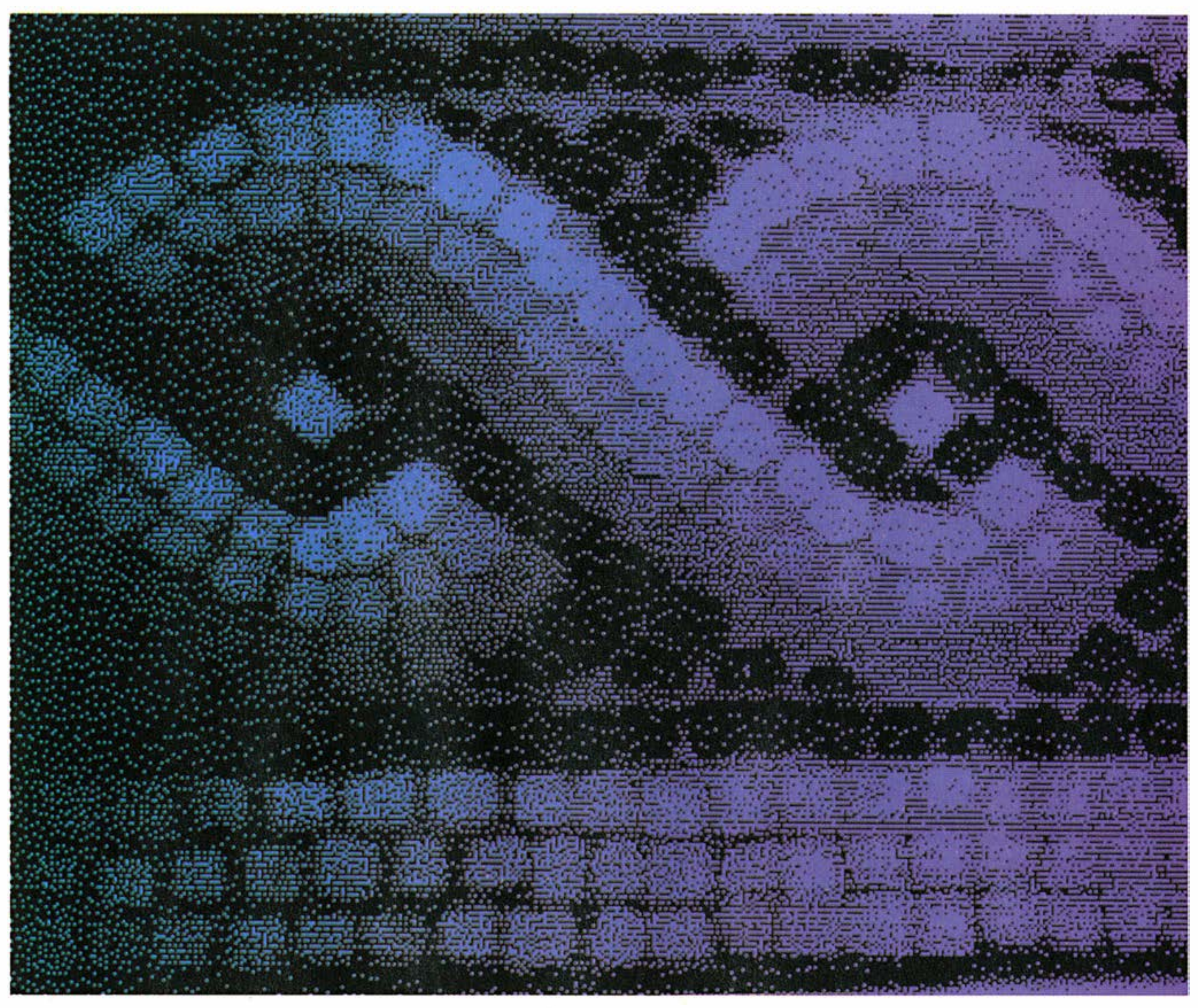
THE POLYMERASE CHAIN REACTION. Dirigido por Kary B. Mullis, François Ferré y Richard A. Gibbs. Birkhäuser; Boston, 1994.

DNA FINGERPRINTING, por M. Krawczak y J. Schmidtke. Bios Scientific Publishers; Oxford, 1994.

ANCIENT DNA. RECOVERY AND ANALYSIS OF GENETIC MATERIAL FROM PALEONTOLOGICAL, ARCHAEOLOGICAL, MUSEUM, MEDICAL, AND FORENSIC SPECIMENS. Dirigido por Bernd Hermann y Susanne Hummel. Springer Verlag; Heidelberg, 1994.

THE USES OF LIFE. A HISTORY OF BIOTECHNOLOGY, por Robert Bud. Cambridge University Press; Cambridge, 1993.

La síntesis de ADN, la síntesis de proteínas y la regulación de la actividad génica son los pilares de la vida de la célula. Al adentrarnos en el estudio del depósito y la transmi-



Fragmento de un mosaico de Pompeya que muestra una sorprendente coincidencia con la estructura del ADN, descubierta ésta dos mil años después de la teselación de aquél. Comentando esa similaridad, escribe Kary B. Mullis, "todo parece indicar, sin embargo, que los romanos creían que los átomos eran cuadrados"

sión de la información (del ADN y del ARN) no debemos perder de vista el horizonte celular en que se desenvuelve. Como no lo pierde Robert Schleif en *Genetics and Molecular Biology*, un clásico redivivo, donde se hallan, además, pormenorizados los fundamentos en que se apoyan los libros que le acompañan en la reseña.

Tres son los organismos cuya investigación ha catapultado la genética hasta su estado actual: una bacteria (*Escherichia coli*), una levadura (*Saccharomyces cerevisiae*) y un díptero (*Drosophila melanogaster*). *E. coli*, procariota de crecimiento rápido, mide entre uno y tres micrometros de longitud por 0,75 micrometros de diámetro; contiene unos 2×10^{-13} gramos de proteínas, 2×10^{-14} gramos de ARN y 6×10^{-15} gramos de ADN. La bacteria se protege con una doble membrana y una capa intermedia de peptidoglicano. En el citoplasma, una solución proteica de unos 200 miligramos por mililitro, reside la mayoría de los 2000 tipos de proteínas que posee la enterobacteria, así como el ADN y los 10.000 ribosomas del procariota; los ribosomas constan de un tercio de proteína y los dos tercios restante de ARN.

La célula eucariota multiplica por mil el volumen de la célula bacteriana: su diámetro ronda los 10 micrometros. Igual que las bacterias, las células eucariotas contienen membranas, proteínas citoplasmáticas, ADN y ribosomas; se distinguen, sin embargo, entre otros factores, por la presencia de proteínas estructurales que forman redes. El ADN de los eucariotas permanece confinado dentro de una membrana nuclear; sólo proteínas pequeñas pueden atravesarla; las de tamaño mayor y el ARN penetran en el núcleo a través de poros especiales. Constituye el ADN un esqueleto regular de 2′-desoxirribosa unida por enlaces fosfodiéster 3′-5′. La información portada por la molécula está especificada por las bases engarzadas en las desoxirribosas. Se emplean cuatro bases: dos purinas (adenina y guanina) y dos pirimidinas (citosina y timina). La estructura química del ARN es similar a la del ADN, con la salvedad de que en el esqueleto del ARN hay ribosa y falta el grupo metilo de la timina, sustituyéndose ésta por uracilo.

Watson y Crick dedujeron la estructura básica del ADN a partir de datos obtenidos por difracción de rayos X, de las estructuras de las bases y de los descubrimientos de Chargaff (según éste, la fracción molar de guanina equivale a la de cito-

sina y la fracción molar de adenina equivale a la de timina). La estructura de Watson-Crick es un par de cadenas de ADN, antiparalelas y orientadas en sentido opuesto, que tejen una hélice dextrógira. Los pares de bases A-T y G-C descansan en el interior de la hélice, en tanto que caen fuera los grupos fosfato. Aunque suelen reconocerse distintos tipos de hélice, los estudios de difracción acaban de encontrar una notable variación de una base a la siguiente. El calentamiento del ADN en una solución de cloruro sódico rompe los enlaces de hidrógeno entre los pares de bases A-T y G-C. No todos los enlaces se rompen a la misma temperatura, y el ADN muestra una zona de transición entre la hélice entera y las bases sueltas. Estas pueden detectarse por su elevada absorbancia en la región ultravioleta del espectro. El ADN desnaturalizado puede reconstituirse y formar la doble cadena *in vitro*. Pero esa capacidad para recomponer secuencias autocomplementarias no es exclusiva del ADN. *In vitro* pueden incluso formarse híbridos ARN-ADN o dúplices ARN-ARN.

Los esqueletos de fosfato confieren a las moléculas de ADN y ARN una carga uniforme por longitud unidad. Así, en electroforesis en geles de poliacrilamida o agarosa, las moléculas migrarán. Las fuerzas de rozamiento o retardo que los geles ejercen en las moléculas migrantes aumentan con la longitud de ADN o ARN, de suerte que, cuanto mayor sea la molécula, más lenta procede a través del gel. Tras la electroforesis, pueden conocerse las ubicaciones de fragmentos específicos de ADN mediante tinción (con bromuro de etidio) o por autorradiografía.

El ADN de las células eucariotas rodea los nucleosomas; éstos forman solenoides, que, a su vez, se retuercen sobre sí mismos en una estructura mayor. Pese a tan compacto empaquetamiento, el ADN debe permanecer accesible a las proteínas reguladoras, a la polimerasa de ARN para la transcripción, a las enzimas reparadoras del ADN y a cualquier otra proteína que precise el contacto con el ADN. Para localizar nucleosomas se emplea la hibridación Southern, técnica de transferencia muy versátil de la genética molecular que combina la electroforesis y la hibridación ADN-ADN o ARN-ADN; la técnica de Southern permite determinar la sensibilidad a la rotura por una ADNasa en la región de cualquier gen que nos interese.

La supervivencia del cromosoma requiere tres propiedades básicas: replicación, segregación fundada en la replicación del ADN y división celular, y replicación y protección de los extremos del cromosoma. En los cromosomas existen múltiples orígenes de la replicación, orígenes que reciben el nombre de secuencias de replicación, porque pueden clonarse en el ADN que se replicará por sí mismo en otras células; pero ese ADN no se parte propiamente en células hijas porque carece de las señales necesarias para la segregación. La porción responsable de la segregación de los cromosomas en células hijas es el centrómero. A medida que la célula se divide, los microtúbulos confinan los centrómeros en las dos células hijas.

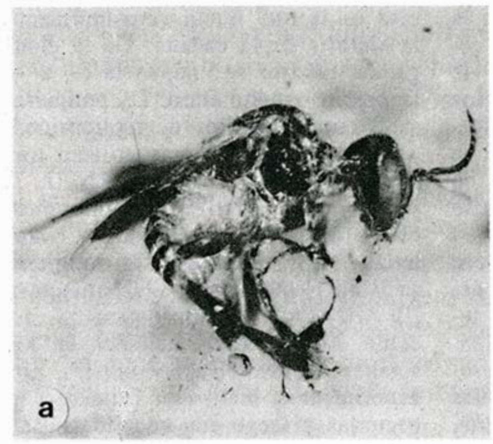
Un error sin corregir en la replicación del ADN es para siempre. Afecta a cada descendiente en que se exprese el gen alterado. De ahí la importancia de la exquisita precisión con que ha evolucionado el mecanismo de síntesis del ácido desoxirribonucleico. Para mantener la precisión se hace obligado el compulsar el texto varias veces y corregir cualquier equivocación. En la replicación del ADN, la criba del error en un nucleótido recién incorporado podría ocurrir antes de que se agregara el nucleótido siguiente; pero podría también realizarse más tarde. Se dan ambas opciones. Las dos cadenas del ADN se replican casi simultáneamente, si bien su polaridad opuesta exige que, conforme la horquilla de replicación avanza secuencia abajo, una cadena crezca en dirección 3′ a 5′ y la otra en dirección 5′ a 3′. Si el nucleótido recién incorporado en la cadena en enlongación no se empareja correctamente con la base de la cadena complementaria, debe excindirse el nucleótido incorrecto. En última instancia, esa tarea editora, atenta a los fallos del lenguaje genético, procurará que la secuencia termine igual que antes de la adición del nucleótido incorrectamente emparejado.

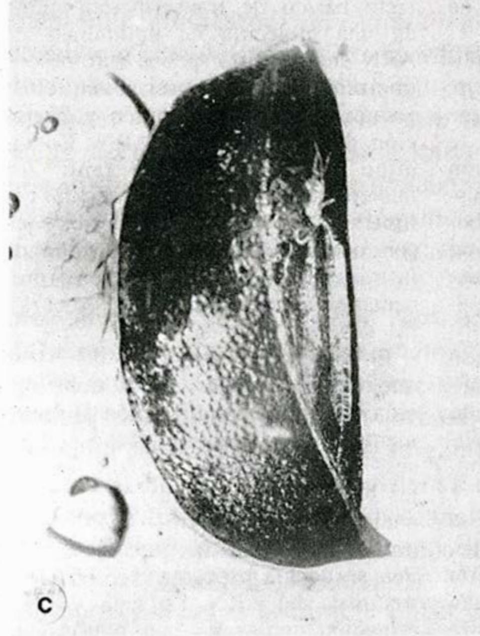
La ADN polimerasa, enzima decisiva en estos procesos de síntesis, debe permanecer unida al complejo de la cadena molde y la cadena en elongación. En las bacterias, la ADN pol III es la principal enzima implicada en la replicación. La pol I cubre los hiatos en la síntesis de la cadena y presta su apoyo en la reparación del ADN dañado. En efecto, ambas, la ADN pol I y la ADN pol III, poseen la actividad exonucleasa 3′ a 5′ que se necesita para la corrección de estilo de nucleótidos mal incorporados. Las células eucariotas poseen también varias polimerasas de ADN: α, β, γ, δ y ε. La α-polimerasa es la que inicia verosímilmente la síntesis de la cadena. De la elongación ulterior se ocupan la δ-, ε- y la propia α-polimerasa. La polimerasa β corrige errores de replicación.

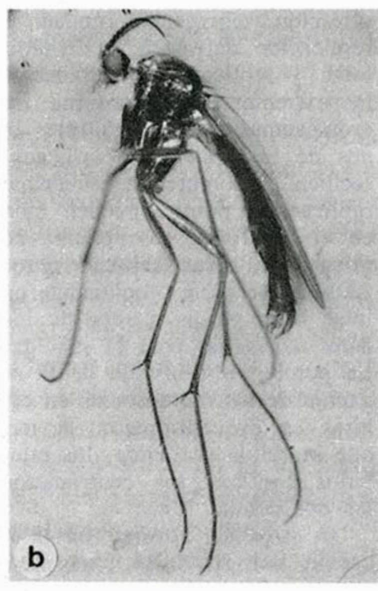
Las células eucariotas utilizan tres tipos de polimerasas de ARN (I, II y III) para sintetizar las diferentes clases de ARN: la polimerasa I sintetiza ARN ribosómico, la polimerasa II ARN mensajero y la polimerasa III ARNt y ARN ribosómico 5S, así como algunos ARN hallados en los somites cirujanos ("splicesomes"). Los experimentos, realizados primero con bacterias y luego con eucariotas, han puesto de manifiesto las fases del ciclo básico de transcripción: unión de una molécula de polimerasa de ARN en un sitio especial (promotor), iniciación de la transcripción, elongación ulterior, terminación y liberación de polimerasa de ARN. El término promotor designa tanto los nucleótidos a los que se une la polimerasa de ARN como otros que son necesarios para la iniciación de la transcripción. Los promotores procedentes de distintos genes procariotas difieren entre sí.

El proceso de elongación del ARN, que ocurre después de la transcripción, abarca la modificación por segmentación, adición de grupos y bases, corte, empalme y edición. En las bacterias, el mecanismo de terminación de la transcripción utiliza la polimerasa de ARN y requiere sólo una secuencia especial cerca del extremo 3′ del ARN. La señal de terminación consta de una región rica en bases GC que pueden formar un bucle estrechamente seguido por una secuencia de uracilos. Esta clase de terminación funciona sin necesidad de factores proteicos auxiliares. En los eucariotas, la terminación suele darse en una secuencia de uracilos también, pero no se deja ver el bucle. Una segunda clase de terminadores proteicos son los que requieren la presencia de la proteína rho.

Gen es un conjunto de nucleótidos que especifica la secuencia de un ARN o una proteína. Su mutación es, pues, una alteración hederable de la normalidad. Los términos "ingeniería genética" y "ADN recombinante" se refieren a las técnicas que cortan y engarzan la molécula de ADN. Permiten así aislar de un organismo un fragmento de ADN, unirlo a otros fragmentos de ADN e introducir el resultado en una bacteria u otro organismo. De ese modo pueden fabricarse muchas copias idénticas (clones). En su ejecución, la ingeniería genética recorre varias etapas. Se aísla el ADN a estudiar, libre de con-







Organismos representativos de ámbar dominicano del que se han extraído muestras de ADN: una abeja de la especie Proplebeia dominicana *(*a*), un díptero (*b*) y una hoja de la leguminosa* Hymenaea protera *(*c*)*

taminantes; se corta por puntos específicos y se obtienen fragmentos que contienen genes o partes de genes; se juntan los fragmentos y se forman moléculas híbridas de ADN; los vectores (plásmidos o fagos), unidos a los fragmentos, penetran en las células. Las enzimas de restricción aportan la herramienta necesaria para cortar fragmentos de ADN. Su exquisita especificidad permite una gran selectividad. A esas manipulaciones genéticas fundamentales hay que sumar ahora técnicas refinadas de clonación, reacción en cadena de la polimerasa, cartografía cromosómica, métodos de alta secuenciación y otros.

Cuenta James D. Watson en el prólogo a *The Polymerase Chain Reaction* que Francis Crick quería introducir en su artículo fundacional de marzo de 1953, sobre la estructura helicoidal de la molécula de ADN, un largo excursus en torno a las implicaciones de dicha molécula. Deseo no compartido del todo, dada la escasez de pruebas, por Watson. En el segundo artículo de *Nature*, de mayo del mismo año, escribieron ya que cada cadena podía actuar "...a la manera de molde para la formación sobre sí misma de una nueva cadena compañera, de suerte que con el tiempo se obtuvieran dos pares de cadenas donde sólo había una", y, además, que "...la secuencia de pares de bases se habría duplicado con exacta precisión". Estaba plantada la semilla que habría de fructificar en la técnica ideada por Kary Mullis a mediados de los ochenta.

La reacción en cadena de la polimerasa (RCP) ofrece tal grado de sensibilidad, que puede detectar una sola molécula de una secuencia específica, incluso en presencia de hasta 10^6 o más secuencias. En un día o dos nos deletrea un segmento de varios miles de nucleótidos a partir de una muestra ínfima. Con ella no se necesita clonación. Viene a ser una suerte de plantilla capaz de multiplicar ("amplificar") el ADN. Para acometer la multiplicación, se comienza por desnaturalizar la muestra, se hibridan luego dos cebadores oligonucleotídicos con el ADN, se elonga la molécula con polimerasa y se itera el ciclo muchas veces. Los dos cebadores oligonucleotídicos deben ser complementarios con cadenas opuestas de ADN. El producto de elongación cebado por un oligonucleótido más el molde se convierten en moldes para la siguiente tanda de síntesis. A resultas de ello, cada tanda de síntesis dobla el número de moléculas de ADN resultantes. Aislada de un termófilo, *Thermus aquaticus*, se usa una polimerasa que resiste la incubación a 95 grados para desnaturalizar el ADN.

La RCP se está empleando en la cartografía y secuenciación de genomas de organismos muy diversos: *Drosophila*, *C. elegans*, *Arabidopsis*, mamíferos y, muy especialmente, el hombre. En medicina, se aprovecha de esta técnica el diagnóstico oncológico. En el foro su incursión facilita enormemente la labor de la justicia. De su capacidad para la identificación de las personas se ocupan M. Krawczak y J. Schmidtke en *DNA Fingerprinting*. El enfoque tradicional en las tareas de individualización consistía en la caracterización de un número variable de repeticiones en tandem mediante análisis del polimorfismo de la longitud del fragmento de restricción a través de la técnica de hibridación de Southern. Aunque se trata de un enfoque valioso y fiable en la prueba de paternidad y forense, tiene sus limitaciones: se requiere cantidad suficiente de ADN de elevado peso molecular, se necesitan sondas marcadas con isótopos para obtener un nivel de detección altamente sensible, el análisis es laborioso y lentísimo y, por último, la técnica en cuestión no puede resolver inequívocamente los alelos de los loci repetidos en tandem. La RCP ofrece, por contra, una mayor sensibilidad y especifidad, comporta menos tiempo y trabajo y no requiere ningún isótopo radiactivo.

En la identificación y clonación de ADN procedente de restos óseos antiquísimos trabajaban desde hace casi un decenio, y sin éxito hasta la aparición de la RCP, Bernd Herrmann y Susanne Hummel, directores de *Ancient DNA. Recovery and Analysis of Genetic Material from Paleontological, Archaeological, Museum, Medical, and Forensic Specimens.* En su rastreo de tejidos antiguos, los expertos emplean cebadores que multiplican segmentos cortísimos (de 100 a 150 pares de bases) y secuencian los productos de RCP para verificar el origen del ADN. La arqueología molecular puede extenderse al ADN "fósil", es decir, procedente de restos que tienen millones de años de antigüedad. Tales logros representan el sueño último del biólogo evolucionista: descubrir en el pasado geológico los procesos de mutación

que quedaron fijados en las especies. Ofrece, en efecto, la oportunidad de reconstruir la historia de una población en el tiempo y en el espacio y, por ende, crear árboles evolutivos.

El análisis evolutivo de las secuencias de ADN tiende un puente entre la filogenética y la genética de poblaciones. El ADN antiguo permite el estudio de genotipos, poblaciones y especies extintos, así como comparaciones de poblaciones y especies vivas. El estudio se ha ejemplificado en la historia del ratón canguro *Dipodomys panamintinus*. Se secuenciaron nucleótidos de la región de control mitocondrial a partir de muestras de museos y muestras modernas de *D. panamintinus*. Y se reconstruyeron las relaciones evolutivas entre genotipos, poblaciones y especies a través del estudio del flujo de genes, el tiempo de divergencia, la cuantificación de la variabilidad genética y su tasa de mutación. Se ha trabajado también en muestras foliares de 17 millones de años y en ámbar fósil, cuyo entorno facilita la conservación de los insectos y el polen encerrado.

Contrapunto de esa visión prometedora en ciencia básica que se funda en los avances de la genética son los resultados, decepcionantes para muchos, en progreso industrial y bienestar de la sociedad. Resume esa situación con objetiva distancia Robert Bud en *The Uses of Life. A History of Biotechnology*. Confiábase en que el maridaje entre microbiología y genética en los años setenta supondría un nuevo renacimiento, parecido al que supuso la zimotecnología para la economía europea del siglo pasado, punto de arranque del trabajo de Bud. (En torno a la industria de la fermentación, de la cerveza en particular, surgieron importantes centros de investigación microbiológica y química, que harían de esa manufactura artesanal una técnica científica. Con la revolución consiguiente de métodos, vale decir, con registro y clasificación de cultivos —cuyo interés algunos asimilaban al del elenco de reactivos—, así como el estudio de la fisiología de la fermentación.)

Aunque el término biotecnología lo acuñó en 1917 Karl Ereky, un economista húngaro empeñado en convertir a su país feudal en una nueva Dinamarca, pionera entonces de la transformación agropecuaria de Europa, la ingeniería genética propiamente dicha nació en 1973, cuando Cohen y Boyer descubrieron la técnica de recombinación del ADN. Logro que posibilitaba, en principio, que las bacterias incorporaran genes (y sintetizaran las proteínas correspondientes) de otros organismos a voluntad.

Científicos, industria y gobiernos de Japón, Estados Unidos y Europa soñaron con un nuevo eldorado en la manipulación de los fundamentos de la vida. Se hicieron proyecciones estadísticas que hablaban de mercados de miles de millones de dólares en el lapso de una generación, ganados con las técnicas recombinantes y de hibridación aplicados a la síntesis de proteínas, mejora de semillas, selección de pesticidas, levaduras y transformación de células humanas patológicas. La realidad, sin embargo, obligó a los laboratorios a dedicarse a la fabricación de fármacos rentables y a la concentración empresarial. Para 1988, sólo se habían aprobado cinco proteínas sintetizadas por manipulación genética: insulina, hormona del crecimiento, vacuna contra la hepatitis B, interferón alfa y activador plasminógeno para lisis o disolución de coágulos. Tan exiguo éxito en la obtención de moléculas de interés indujo a pensar en otra línea de investigación, la de domeñar ingenierilmente los procesos celulares implicados. Y ahí estamos. (L. A.)

Ciencia andalusí

Astronomía popular

Kitab al-anwa' wa-l-azmina -al-qawl fi l-suhur. Libro sobre los anwa' y los tiempos —Capítulo sobre los meses— de Ibn Asim (m. 403/1013). Estudio, traducción y edición crítica por Miquel Forcada Nogués. C.S.I.C., Instituto de Cooperación con el Mundo Arabe e Instituto Millás Vallicrosa de Historia de la Ciencia Arabe. Barcelona, 1993.

El tratado sobre los anwa' y los tiempos de Ibn Asim es uno de los raros libros del anwa' que ha llegado completo hasta nuestros días que, en esta reciente edición, se presenta parcialmente editado en árabe (el capítulo referente a los meses) con traducción y estudio a cargo de Miquel Forcada Nogués, de la escuela barcelonesa de historiadores de la ciencia árabe que fundó por los años treinta J. M. Millás Vallicrosa, bien conocida por los lectores de esta revista.

Los libros de anwa' (singular: naw') son compilaciones sobre conocimientos de astronomía y meteorología popular que se remontan a los inicios de la cultura árabe y cuyo contenido ha evolucionado a lo largo de los siglos.

El tratado de Ibn Asim data de las proximidades del año 1000 y está recogido en un único manuscrito del siglo XIII que se conserva en la biblioteca del Topkapi Saray de Estambul. Su autor, Ibn Asim, procede de una noble familia andalusí afincada en Córdoba a quien se atribuyen unas pocas obras, la más importante de las cuales es el presente *Libro de anwa'*.

Los materiales que utilizó Ibn Asim proceden de tratados anteriores, principalmente los de Ibn Qutayba y de Abu Hanifa, que ordena, reescribe y, a veces, adapta a la realidad andalusí de su tiempo. Cronistas contemporáneos aseguran que el libro de Ibn Asim fue muy útil y consultado por las gentes, aunque no fuera tan apreciado por las clases más cultivadas que podían disponer de los originales de los que se sirvió Ibn Asim. La importancia del libro fue creciendo con el paso del tiempo e incluso fue tomado como fuente para la elaboración de tratados posteriores.

El concepto de anwa' relaciona los fenómenos astronómicos observables, como son el orto helíaco y el ocaso acrónico de determinados astros, con las 28 mansiones lunares. Estos anwa' determinan períodos de 13 días que sirven de marco zodiacal para ubicar dichas observaciones repetidas anualmente; a la vez, se circunscribe a cada naw' un tiempo atmosférico o fenómenos meteorológicos que se supone se repetirán año tras año por las mismas fechas y que se piensa dependen directamente de la aparición de determinados astros en el cielo del amanecer. Así, el sistema de los anwa' es una especie de calendario meteorológico que comprende ciclos de 13 días y que sirve de complemento entre el tradicional calendario lunar árabe y el año solar que determina el ciclo de las estaciones.

Los anwa' nacieron de la observación de determinados efectos climáticos y de su aparición y reproducción regular coincidiendo con el orto y ocaso de determinadas estrellas en el firmamento antes del alba. Los primeros compiladores de este saber popular lo sistematizaron e insertaron en el sistema astronómico culto de las mansiones lunares de origen hindú.

Los orígenes, la evolución y la progresiva utilización de dicho sistema de los anwa' en la cultura islámica se explican en la primera parte del estudio de Forcada, en la que aborda también los tratados y tratadistas populares y científicos de anwa' para centrarse seguidamente en la persona de Ibn Asim, su formación,

su vida y sus obras, el entorno cultural en el que vivió y las influencias que tuvo en la redacción de su tratado. Asimismo, Forcada compara el contenido del tratado de Ibn Asim con el célebre "Calendario de Córdoba" y reseña la repercusión de su obra en autores posteriores.

Una extensa parte del estudio está dedicada a la disección del contenido del tratado de Ibn Asim que comienza por definir los conceptos astronómicos que aparecen: la descripción del cielo *—todo aquello que está sobre ti y te cubre—*, la esfera celeste *—lugar por donde circulan las estrellas—*, el espacio, el horizonte, el polo, los signos del zodíaco, las mansiones lunares, el ocultamiento y visibilidad de la Luna, el significado de los anwa', los ocasos y los ortos de los astros y los siete planetas.

Siguen a continuación una serie de capítulos dedicados a los meses del calendario solar con la explicación de las mansiones lunares que comprende cada uno, el aspecto de la bóveda celeste y los fenómenos astronómicos observables durante su vigencia, así como las manifestaciones naturales asociadas de índole meteorológica, cronológica, botánica, zoológica y también la relativa a la medicina, farmacología y navegación.

Es de destacar el léxico astronómico empleado por Ibn Asim y la etimología de cada nombre reseñado donde se ofrece la versión tradicional propia de la herencia preislámica y la adoptada de la cultura clásica.

Siguen otros capítulos sobre distancias angulares entre estrellas, las Pléyades, estrellas de astrolabio, estrellas rojas, la Osa Mayor, la Osa Menor y la Vía Láctea *—madre de las estrellas*. Ibn Asim reserva varios capítulos finales de su obra a la astronomía sagrada, en los que se trata de la determinación de la hora de las oraciones por la longitud de las sombras, el momento de la ruptura del preceptivo ayuno de Ramadán y la determinación de la dirección de La Meca. (E. F.)

Genética

Escuela de Morgan

LORDS OF THE FLY. *DROSOPHILA GENETICS* AND THE EXPERIMENTAL LIFE, por R. E. Kohler. Chicago University Press; Chicago, 1994.

Un libro con este título debería atraer a miles de drosofilistas, como las drosófilas son atraídas por



T. H. Morgan en su laboratorio de trabajo, en la "sala de las moscas" de la Universidad de Columbia

la fruta madura. Cientos de laboratorios de todo el mundo trabajan hoy con *Drosophila melanogaster* como organismo experimental. Este díptero es el organismo multicelular mejor conocido en su genética, en su bioquímica y aun en su ecología y evolución. En gran parte por ello, Robert Kohler (un historiador de la ciencia) ha elegido este organismo experimental como modelo para un estudio histórico-social de las motivaciones que guían a una comunidad de investigadores.

Kohler analiza la empresa científica iniciada por T. H. Morgan en 1910 en la "Fly Room" de Columbia y continuada desde 1928 en el CalTech hasta 1940. Esta es primeramente una empresa humana, de adaptación del investigador a las peculiaridades de la mosca y su domesticación. Es una empresa colectiva, en la que los investigadores trabajan interactivamente aislando problemas, buscando medios, definiendo objetivos. Es una empresa que evoluciona: con indecisos comienzos, entusiasta exploración y diáspora colonizadora —y finalmente ocaso, cuando los intereses de los biólogos se trasladan a organismos más simples como hongos, bacterias y fagos. Hace treinta años muy pocos laboratorios permanecían todavía trabajando en *Drosophila*. La explosión reciente es un renacimiento, con lo que esto supone de recuperación del espléndido material acumulado por los héroes clásicos, los "Señores de la Mosca".

El autor nos muesta que estos cambios de interés o de objetivos se manifiestan ya en la etapa clásica. Se corresponden a evoluciones determinadas por constantes del método científico (elección del problema, su exploración y su agotamiento), búsqueda de alternativas, dificultades técnicas, y, no en postrer lugar, ambición de los investigadores individuales o prestigio del grupo. El éxito o fracaso del grupo depende de las relaciones de interdependencia entre investigadores, sus orígenes sociales, contextos académicos, la financiación y sus ambiciones.

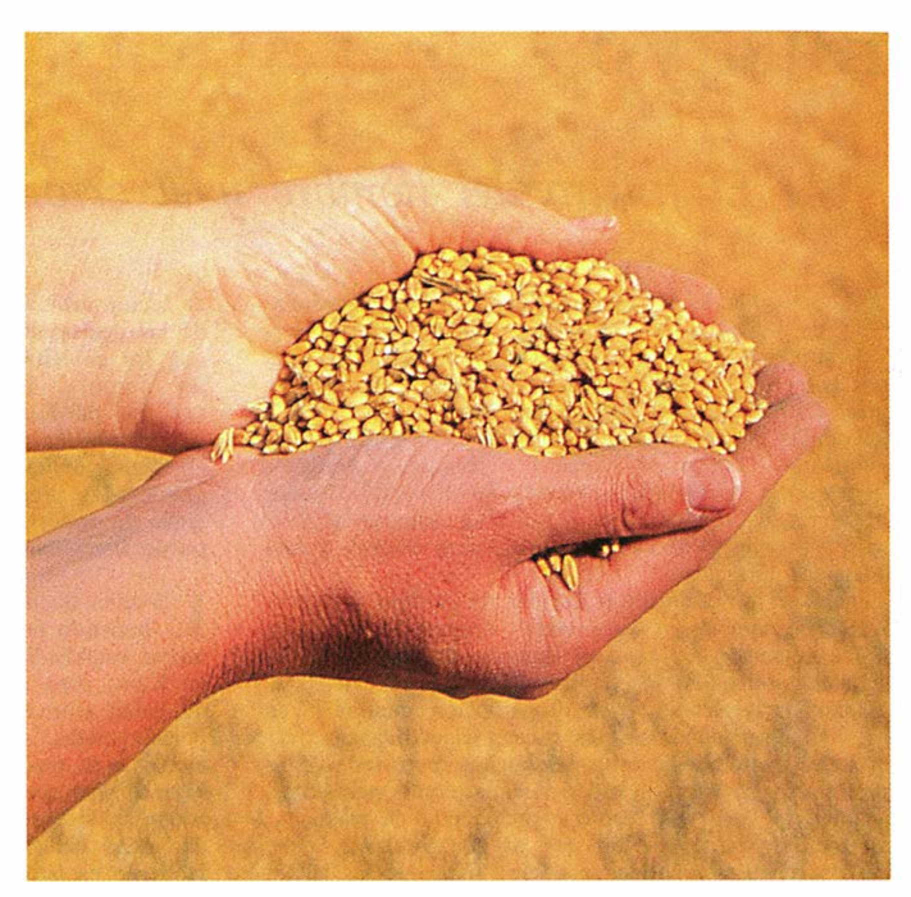
El contenido del libro tiene como urdimbre una relación de los descubrimientos clásicos en la genética de *Drosophila*, realizados por el grupo que organizó y dirigió T. H. Morgan desde su origen hasta su muerte en 1945. Nombres que encontramos en los tratados de texto aparecen en el libro descubriendo mutantes, cartografiándolos en cromosomas, analizando series fenotípicas, comparando cromosomas y luego sus mapas genéticos entre especies de Drosofílidos en Columbia (A. H. Sturtevant, C. Bridges, H. Muller, Ch. Metz), y, después del traslado del grupo al CalTech y con la llegada de visitantes europeos, en la evaluación de mosaicos genéticos y con ello la apertura de la genética fisiológica o bioquímica y de variantes cromosómicos en poblaciones naturales (C. Stern, Th. Dobzhansky, G. W. Beadle y B. Ephrussi).

Pero el libro no es una biografía de sus héroes, ni siquiera un análisis de sus ideas científicas; es una historia del sistema, de la estructura que permitió y condicionó el avance científico. La documentación histórica es excelente. Sus reflexiones sobre cómo el mundo real condiciona el avance científico son originales y sugestivas. Es una lectura apasionante.

La evolución del grupo va asociada a la creación de una ética de ayuda mutua y compromiso de intereses, pero también de conflictos internos, resultado de la convivencia y los celos, así como externos, por la diáspora de sus miembros y aparición de nuevos grupos competitivos en universidades, centros de investigación en Estados Unidos y en Europa.

Creo que el énfasis en un análisis sociológico de la aventura del descubrimiento científico es en sí mismo válido, si se puede extrapolar al de cualquier empresa científica, en cualquier momento histórico. Pero el enfoque social, que busca un determinismo histórico, soslaya la contri-

bución del investigador singular en el proceso de buscar, arriesgarse, descubrir y extender sus investigaciones. Las motivaciones singulares se exponen en las biografías singulares (y las hay excelentes de algunos de los investigadores nombrados), pero estos aspectos quedan diluidos o parecen inoperantes en una visión de estrategias de grupo. Y este es el problema de este libro, si de su lectura queremos sacar una lección de historia para nuestro proceder en la ciencia que nos ha tocado desarrollar hoy día. En los héroes clásicos su motivación era el éxito en la cacería, la satisfación del trabajo intenso, la búsqueda de lo relevante, el reconocimiento y admiración de sus pares. Posiblemente hoy día las motivaciones sean más simples, un puesto de trabajo y una retribución suficiente, una contribución honrada lo menos elitista posible. En ese sentido, este libro de la historia de la genética clásica tiene poco que ofrecer a una juventud que no cree en dioses, en sacrificios y en destinos y ciertamente menos en Olimpos, peor aún, que cree no tener nada que aprender de la historia. (A. G.-B.)



Granos, frutos y otros productos de las flores constituyen la mayor parte del alimento consumido por el hombre. La capacidad de manipular el desarrollo y la forma de las flores puede ayudar a crear algún día cosechas más productivas. No sólo eso, muchas especies vegetales, todavía desconocidas, encierran interés farmacéutico que se perdería con la desaparición de las mismas

Biodiversidad

De moda

BIODIVERSITY, TEMPERATE ECOSYSTEMS AND GLOBAL CHANGE. Dirigido por Timothy J. B. Boyle y Christopher E. B. Boyle. Springer Verlag; Berlín, 1994.

El éxito de los conceptos de biodiversidad y cambio global, junto con una mayor facilidad para organizar reuniones internacionales, han generado la proliferación de volúmenes, que en parte suscitan la impresión del *déjà vu*, pero que siempre contienen algunas perlas, más o menos ocultas. Este reúne los materiales preparados en un Taller Avanzado de Investigación, que se reunió del 15 al 19 de agosto de 1993, en Montebello, Canadá.

El libro consta de una introducción y de 24 capítulos, de los que seis son de más de un autor; en su mayoría, los autores son conocidos por haber publicado anteriormene sobre estos temas. Al restringir el estudio a los ecosistemas de la región templada, en el sentido de Köppen, definida así porque durante cuatro o más meses la temperatura excede de 10° C, la diversidad es fluctuante y nunca llega a ser muy alta, en relación con la periodicidad asociada al ciclo ambiental anual. De ese modo escapamos a la reiterada referencia a los sistemas de alta diversidad ("selva amazónica"), que son los que gozan de mayor popularidad y nos circunscribimos a ecosistemas que son realmente importantes en la región templada, donde si es cierto que hay muchos problemas, tal vez son mayores las facilidades de manejo.

En ello reside la novedad o la ventaja de este volumen que comenta resultados recientes o nuevos en el estudio de animales que viven próximos al substrato (coleópteros, M. Loreau, A. George; artrópodos del suelo, R. Pissolotto; aves, A. J. Hansen y colabs.). Las referencias a la vegetación de los bosques son numerosas y reflejan el carácter reconstruido y humanizado de la mayoría de los bosques templados. Se hace notar la pérdida de diversidad de la flora de hongos y el carácter de inmigrantes recientes de muchos de los habitantes en los sistemas costeros reconstruidos de los Países Bajos (W. J. Wolff y colaboradores); en este texto se cuela el error de admitir el exotismo de *Asterionella gracialis* (*japonica*).

Aunque una gran parte de los datos introducidos o recopilados en el texto no encierran sorpresa o novedad, hay piezas de investigación y resúmenes que resultan útiles. Las estrategias de conservación que pueden proponerse para el Canadá difícilmente serían viables en otros países con una mayor presión de la población y pueden ser, en parte, discutibles.

Como es costumbre, no se debaten los conflictos reales que aparecen en las relaciones entre poblaciones humanas que adoptan estrategias diferentes: las ricas dispuestas a limitar la tasa de aumento poblacional y las pobres que apenas cuidan de ello. Si no hay una apreciación real de este tipo de problemas, es muy difícil esperar que se unifiquen voluntades y se pueda aprobar una estrategia en relación con un programa de conservación global. La presentación es excelente, con un índice alfabético particularmente detallado. (R. M.)

Ensayo

Antonio R. Damasio

El futuro del hombre y el error de Descartes

A comienzos de los años cincuenta, en una apasionada alocución inspirada por la amenaza de la destrucción nuclear, William Faulkner echó en cara a sus colegas los escritores el haber "olvidado los problemas del corazón humano en conflicto consigo mismo". Y les pidió que no dieran cabida en sus obras "a cosa alguna que no fuesen las viejas convicciones y verdades del corazón, a las antiguas verdades universales —el amor y el honor, la piedad y el orgullo, la compasión y el sacrificio—, sin las cuales toda historia es efímera y está condenada al fracaso..."

Aunque la grave amenaza nuclear ha perdido dramatismo, nos enfrentamos a otros claros riesgos: la población mundial aumenta; el aire, el agua y los alimentos se siguen contaminando; los patrones éticos y educativos se degradan; la violencia y la drogadicción continúan en auge. Son muy numerosas las causas específicas que actúan en el trasfondo de estos fenómenos, pero todas ellas dimanan de lo irracional de muchos comportamientos, irracionalidad que se está propagando como la peste.

Yo siempre he entendido que lo que Faulkner quiso decir es que la racionalidad necesaria para que los humanos superen esos retos y perduren ha de estar imbuida de los sentimientos y emociones que brotan de lo más profundo de nuestro ser. Esta opinión se me hace atractiva, porque la emoción es un factor integrante de todo proceso discursivo. Sospecho, inclusive, que de lo que adolece hoy la humanidad no es de falta de competencia lógica, sino más bien de una falta de emociones que informen el despliegue de la lógica.

¿Qué pruebas puedo aducir en apoyo de estos enunciados contraintuitivos? Pues las que se obtienen estudiando a individuos que, habiendo actuado antes racionalmente, por sufrir lesiones neurológicas en determinados sistemas del cerebro perdieron la capacidad de tomar decisiones racionales y, junto con ella, la de procesar normalmente las emociones. Sus instrumentos de racionalidad aún pueden recuperarse; el conocimiento del mundo en el que han de obrar sigue siendo asequible; y su capacidad para abordar la lógica de un problema permanece intacta. Sin embargo, muchas de sus decisiones personales y sociales son irracionales. Yo he sugerido que lo que en tales casos ocurre es que al delicado mecanismo del razonar no le afectan ya los influjos emocionales que deberían afectarle.

Este tipo de pacientes suele tener deterioradas unas áreas muy precisas de las regiones cerebrales frontal, temporal y parietal derecha, pero hay otras condiciones cuya causa neurológica todavía no ha sido identificada y que presentan características similares en muchos aspectos. Los sociópatas de que nos hablan los noticiarios son individuos inteligentes y lógicos que, no obstante, carecen de emotividad normal. Su comportamiento irracional es destructivo para sí mismos y para la sociedad.

Sin duda, puede decirse que la falta de emoción es por lo menos tan perniciosa para la racionalidad como el exceso de emoción. Ciertamente no parece que a la razón se le fueran a seguir muchas ganancias del actuar sin el influjo de la emoción. Al contrario, es probable que ésta nos ayude a razonar bien, sobre todo en los asuntos personales y en los sociales, orientándonos a veces hacia el sector más ventajoso para nosotros dentro del campo de las decisiones posibles. En breve, no estoy sugiriendo que las emociones sustituyan a la razón o que decidan por nosotros. Tampoco estoy negando que la emoción excesiva pueda generar irracionalidad. Lo único que digo es que nuevos datos neurológicos parecen indicar que la absoluta carencia de emociones es un problema aún más grave. Muy probablemente, la emoción constituye el sistema de soporte sin el que el edificio de la razón no puede funcionar bien.

La idea de que el bastión de la lógica no debe ser invadido por la emoción y el sentimiento está firmemente establecida; se la encuentra tanto en Platón como en Kant; pero tal vez no hubiese sobrevivido nunca de no haber sido expresada tan vigorosamente como lo fue por Descartes, que se complació en separar de las emociones a la razón y apartó a ésta de sus cimientos biológicos. Por supuesto que la escisión cartesiana no es la causa de las patologías que hoy padece la razón, pero sí que debe achacarse a aquélla lo mucho que el mundo moderno ha tardado en reconocer su raigambre emocional. Cuando se concibe a la razón como carente de ascendencia biológica, es bastante fácil pasar por alto el papel que en su funcionamiento desempeñan las emociones, bastante fácil no advertir que nuestras decisiones presuntamente racionales acaso sean manejadas por las emociones que queremos mantener a raya, bastante fácil despreocuparse de las negativas secuelas que puedan tener las experiencias emocionales en que se vive como entretenimiento la violencia, y bastante fácil ignorar el efecto positivo que las emociones bien armonizadas pueden producir en la gestión de los asuntos humanos.

No es probable que la razón comience a existir con el pensamiento y el lenguaje en un denso ambiente cognitivo, sino más bien que se origine de la evolución biológica de un organismo viviente empeñado en sobrevivir. El núcleo cerebral de organismos tan complejos como los nuestros contiene, en efecto, un refinadísimo aparato que capacita para tomar las decisiones relacionadas con el mantenimiento de los procesos vitales. Entre las respuestas de este aparato se incluyen la regulación del medio interno y los impulsos, instintos y sensaciones. Sospecho que la racionalidad depende de la enérgica pasión por la razón que anima a tal aparato.

Es curioso comprobar que Pascal, en el mismo siglo XVII en que se introdujo el dualismo cartesiano, vislumbró ya esta otra idea cuando dijo: "Sobre esta sabiduría del corazón y de los instintos debe basarse la razón misma y echar los fundamentos de todo su discurso." Nosotros estamos empezando a descubrir los pertinentes datos neurobiológicos que, a no mucho tardar, confirmarán la profunda visión pascaliana.

ANTONIO R. DAMASIO ocupa la cátedra M. W. Van Allen y dirige el departamento de neurología en la facultad de medicina de la Universidad estatal de Iowa.